

# Powder Diffraction on proteins

Irene Margiolaki



ESRF

# Presentation

A Challenging Project

Protein Crystallography via powder diffraction

Developments

- Improvements in data quality
- Novel methods for data analysis: Case studies of small proteins
- Cryocooling of protein powder samples

Perspectives

Projects & Collaboration with Industry

# Rietveld Methods....

Use the entire powder pattern on a point by point basis

Bragg Reflections

Peak Intensities

Structure  
Positions  
Occupancies  
Thermal factors

Background

Peak shapes

Peak positions

Composition  
("Phase" analysis)

Non-crystalline components

Microstructure  
Size, Shape, Strain

Cell parameters

The peak overlap problem makes all of this more difficult!

# Protein Crystallography

## TRADITIONAL METHOD:

### Single Crystal Diffraction

- Gives 3D information about crystal structure
- Very small amount of protein required

## HOWEVER

- Can be very difficult to grow large enough single crystals of some proteins
- Very specific crystallisation conditions may not represent the natural environment

## COMPLEMENTARY METHOD:

### Powder Diffraction

- Three dimensional information is collapsed down into 1D – loss of information!
- Larger protein sample required

## HOWEVER

- Polycrystalline powder can often be obtained when a good single crystal cannot
- More crystallisation conditions possible
- Phase mixtures and phase transitions can be observed in-situ

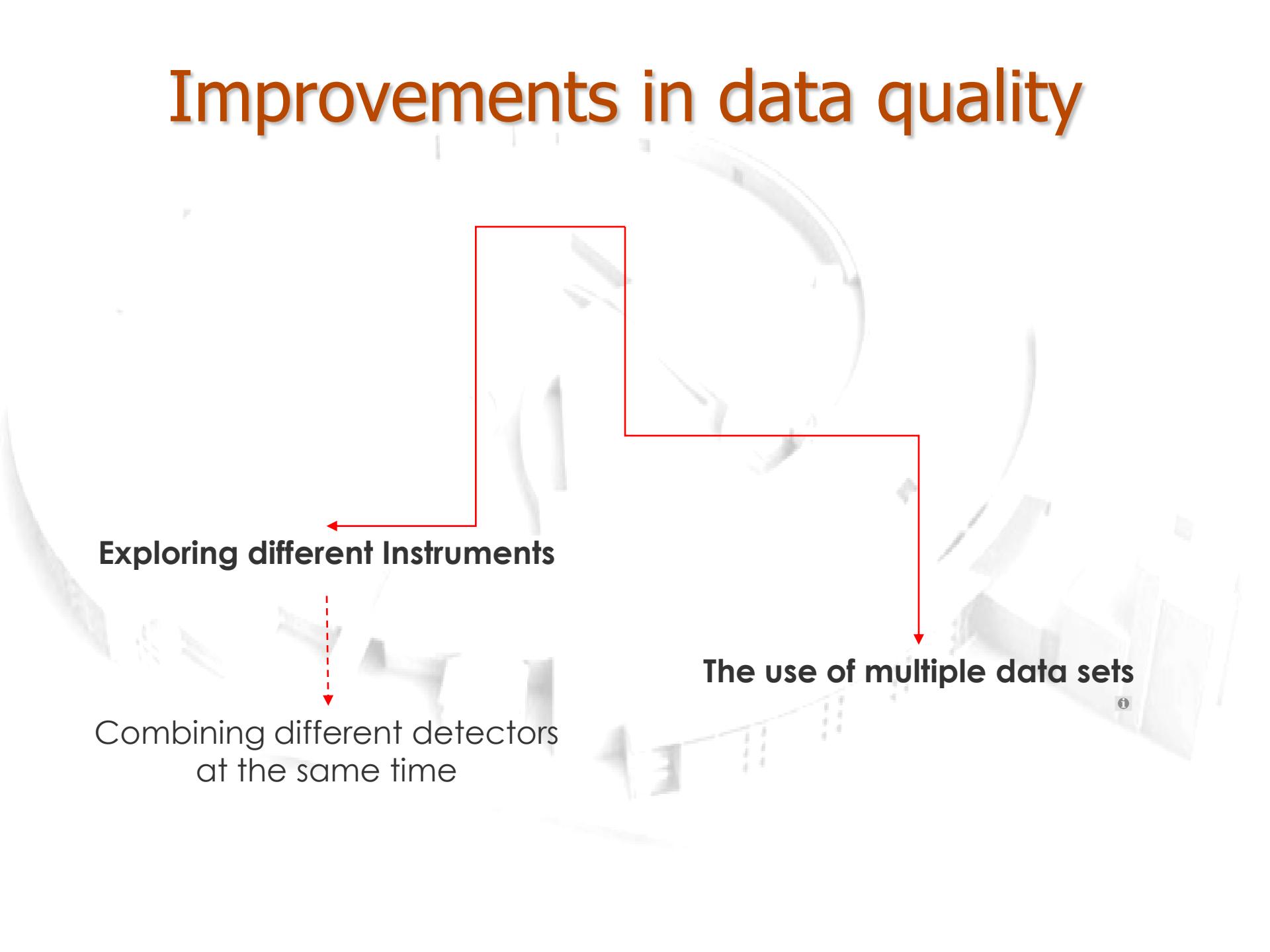
# Proteins Characterised successfully using Powder Data

15 proteins characterised using synchrotron data

41 structural models deposited in the Protein Data Bank (PDB)  
employing powder diffraction

<http://www.rcsb.org/pdb/home/home.do>

# Improvements in data quality

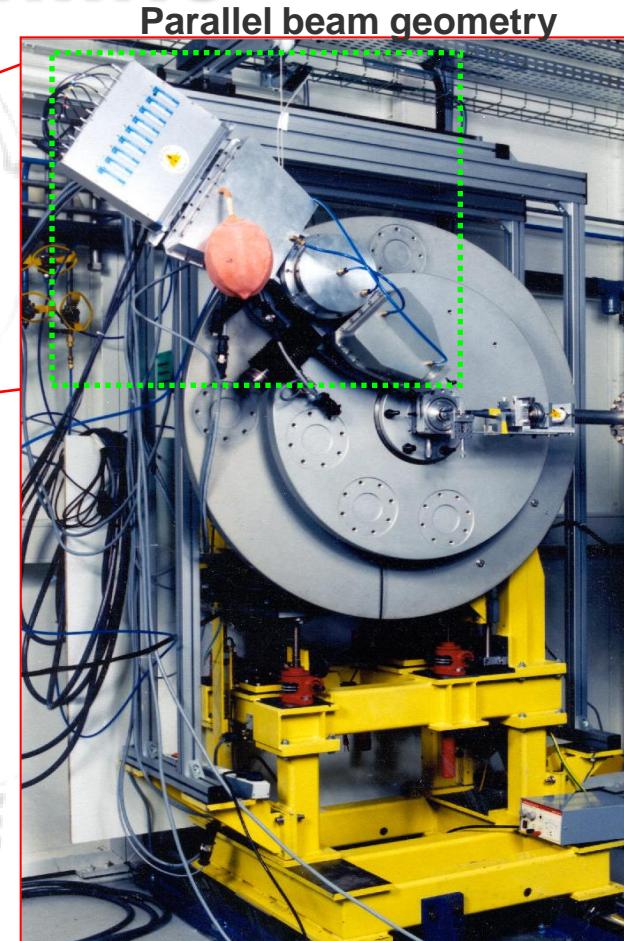
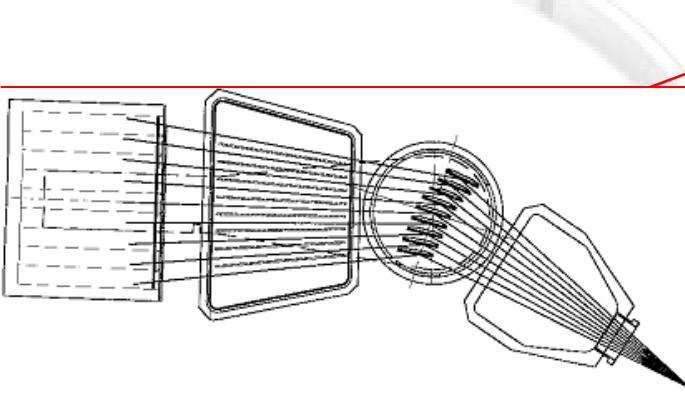
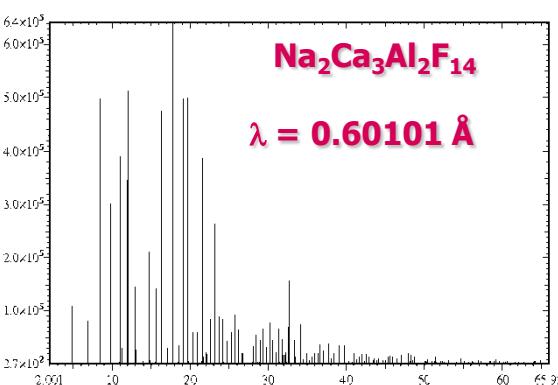


Exploring different Instruments

The use of multiple data sets

Combining different detectors  
at the same time

# ID31: High Resolution Powder diffraction Beamline



## Experimental Conditions:

He- Cryostat for measurements down to 3 K.

"Hot-air blower": Measurements up to 950 °C

Oxford Cryosystems Cryostream (Cold Nitrogen Gas blower) Measurements in the temperature range:  
80 - 500 K

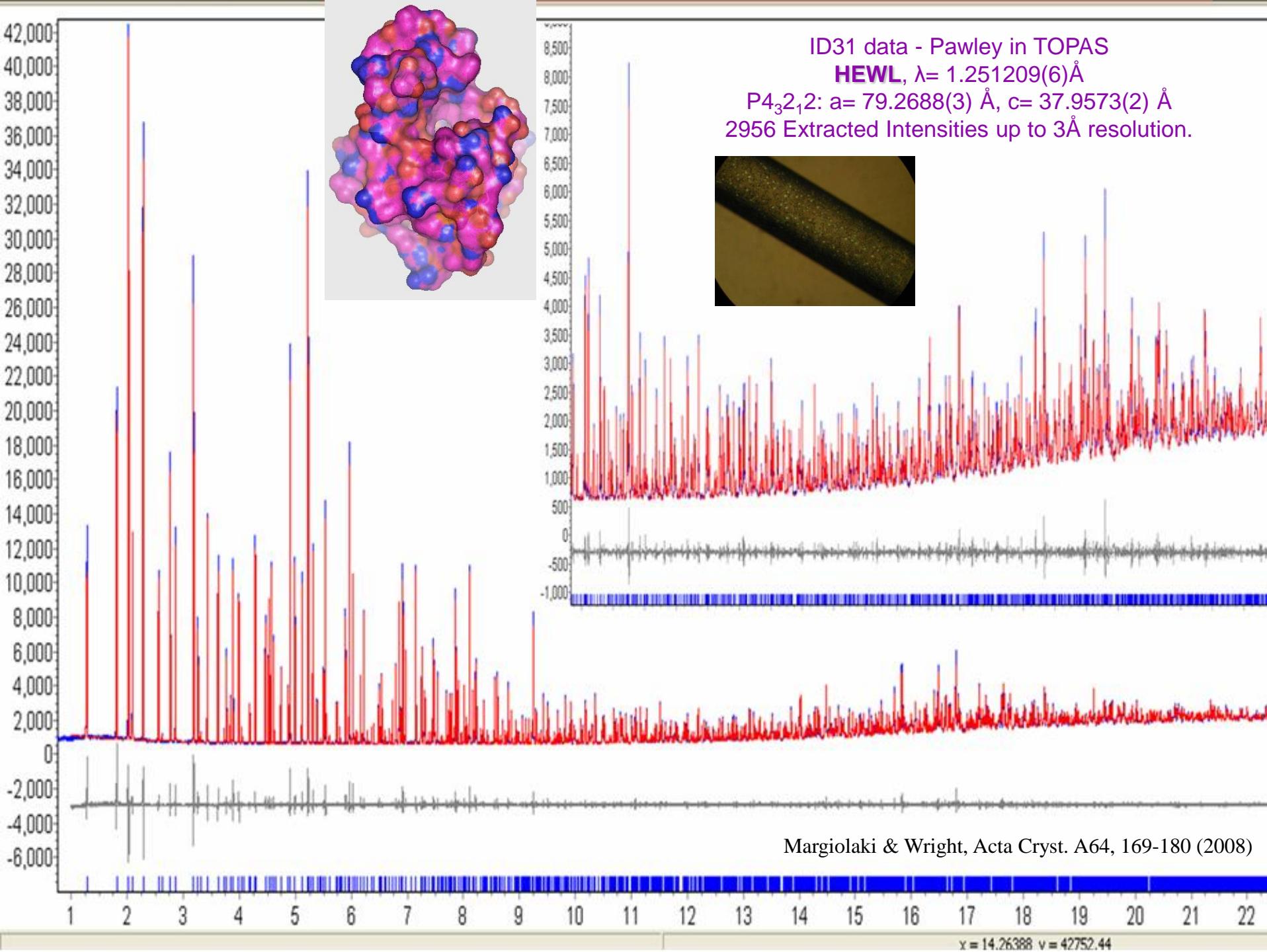
"Mirror Furnace": measurements up to ~1500°C (Pt capillaries)

Translation stages (XYZ) for in situ measurements of residual stress.



Detection of scattered beam using 9 APD detectors and 9 Si(111) analyser crystals

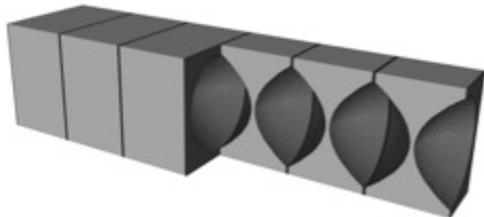
Thanks to J.-L. Hodeau, M. Anne, P. Bordet, A. Prat, CNRS, Grenoble.



# Area detectors available at ID11/ESRF

- Materials Science Beamline

- High energy = small angles
- Low absorption



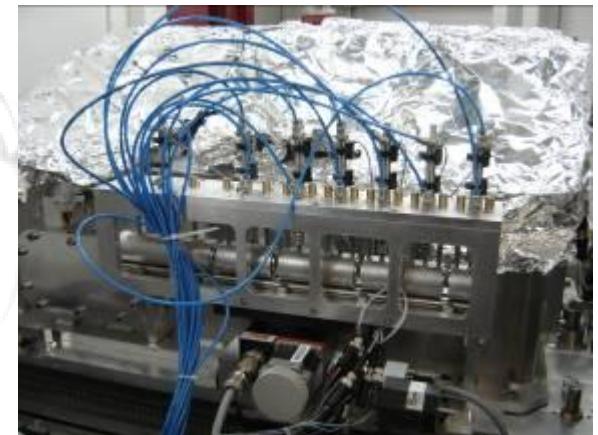
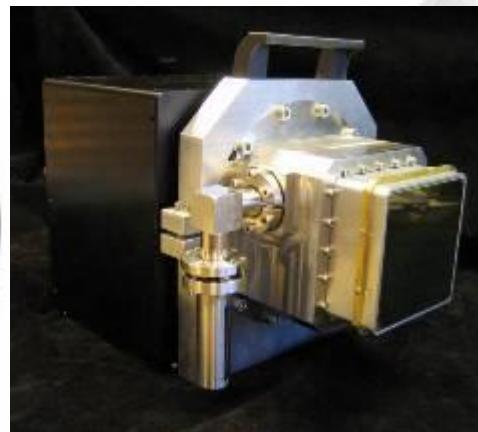
Focus to  
~10 $\mu$ m at  
detector

- Aluminium compound refractive lenses

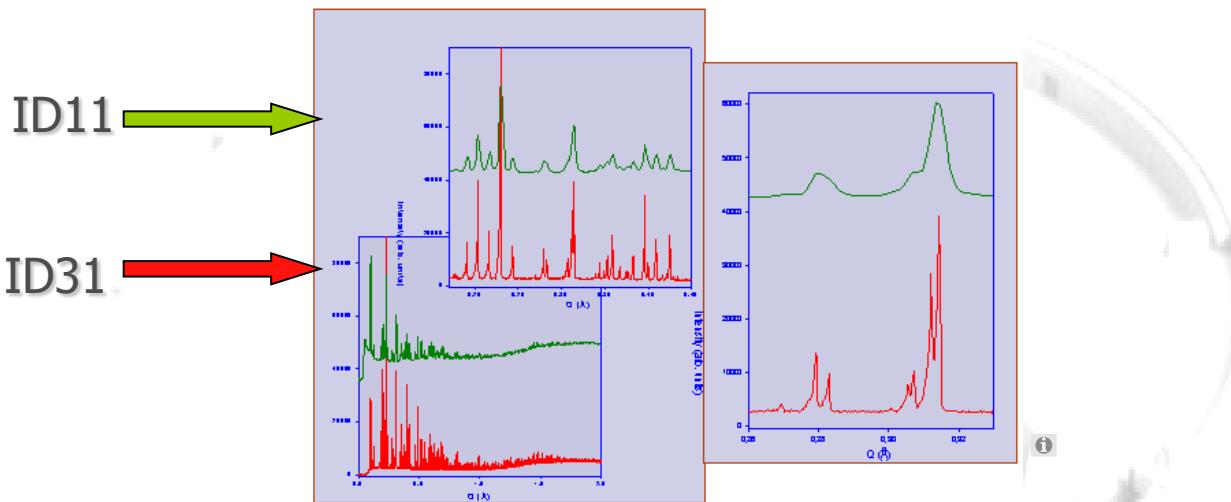
- Focus at detector
- Transfocator: Snigirev/Rossat

- Variety of area detectors

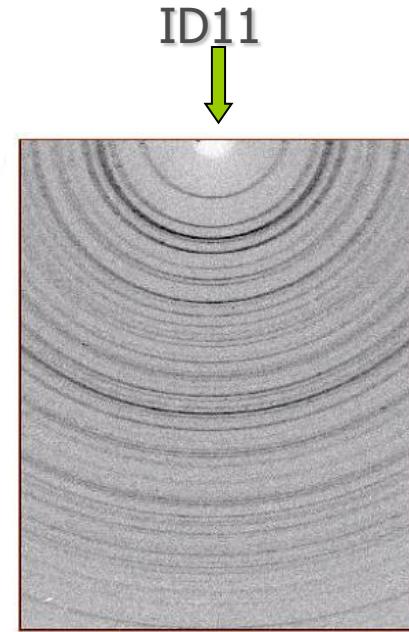
- Bruker smart 6500
- ESRF Frelon
- Trixel / Pixium



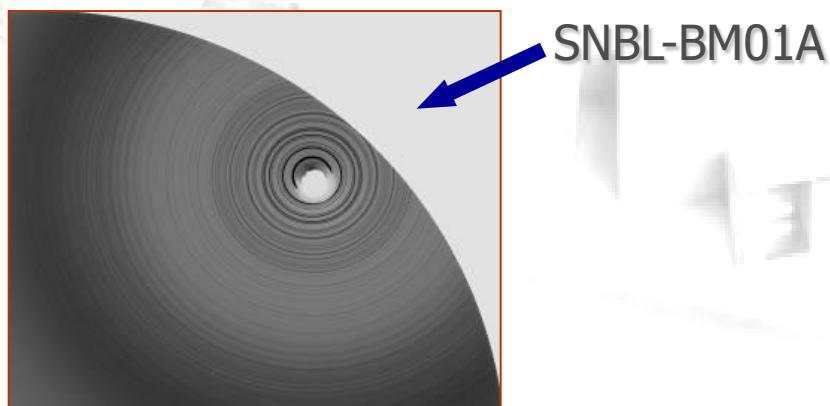
# Exploring different detection systems



HEWL data collected at the analyzer crystal beam line- ID31 (red line), the area detector beam line- ID11 using a CCD camera (green line). Inset: Magnification of the low angle data,



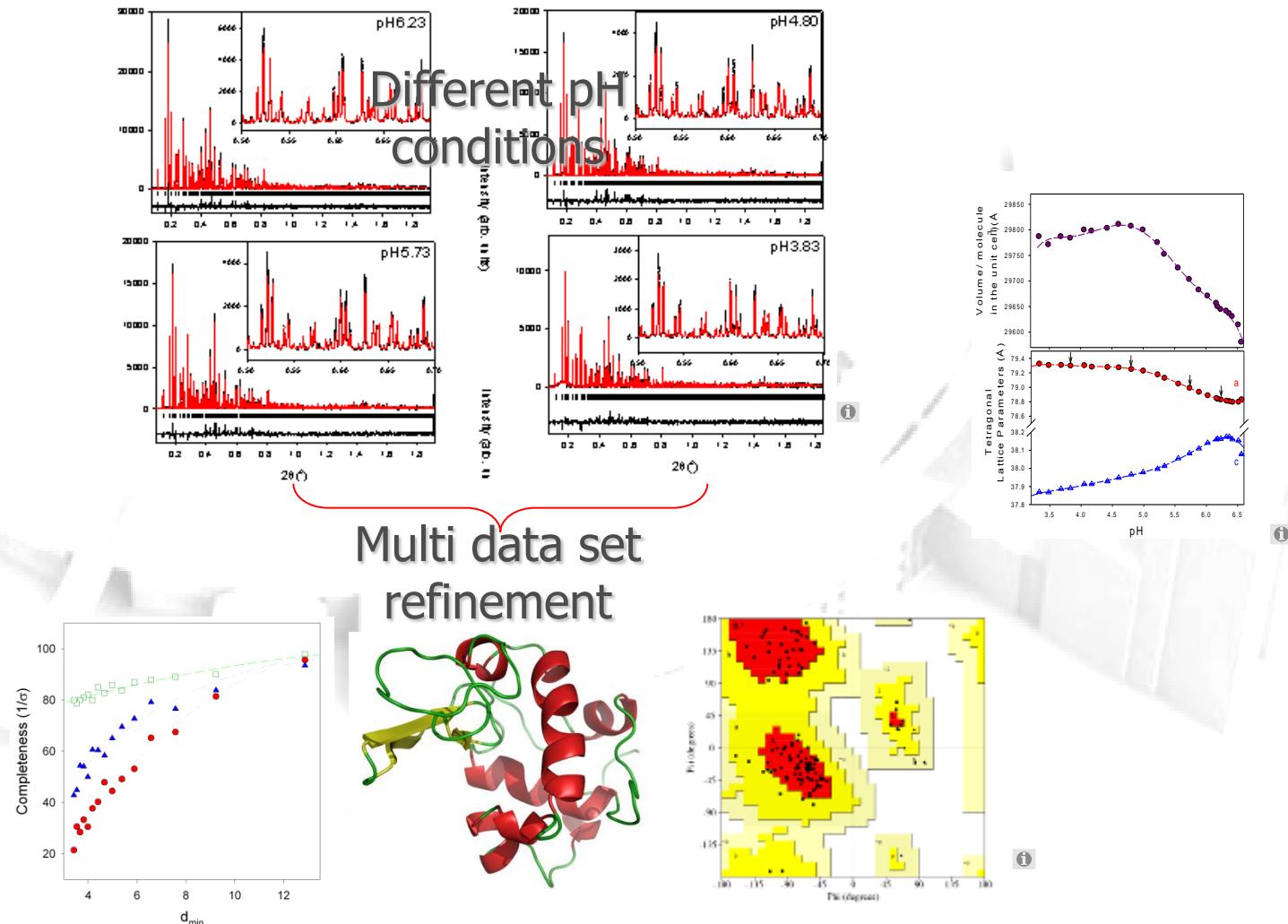
ID11 using a FReLon camera



The area detector station of SNBL- BM01A.

I. Margiolaki, J. P. Wright, A. N. Fitch, G. C. Fox, A. Labrador, R. B. Von Dreele, K. Miura, F. Gozzo, M. Schiltz, C. Besnard, F. Camus, P. Pattison, D. Beckers, T. Degen,  
Z. Kristallogr. Suppl. 26 (2007) 1-13

# The multi-data-set Rietveld refinement for HEWL samples at pH 6.23, 5.73, 4.80, 3.83, crystallised at 277 K.



## Improved methods for Intensity extraction and refinement via the use of multiple profiles

- PRODD, Wright et al. Z. Kristallogr. Suppl. 26 (2007) 27-32
- GSAS, Von Dreele, R. B. (2007). J. Appl. Cryst. 40, 133–143 & Basso et al., Acta Cryst. D61, 1612-1625 (2005)

## Combination of Software designed for single crystal and powder diffraction data

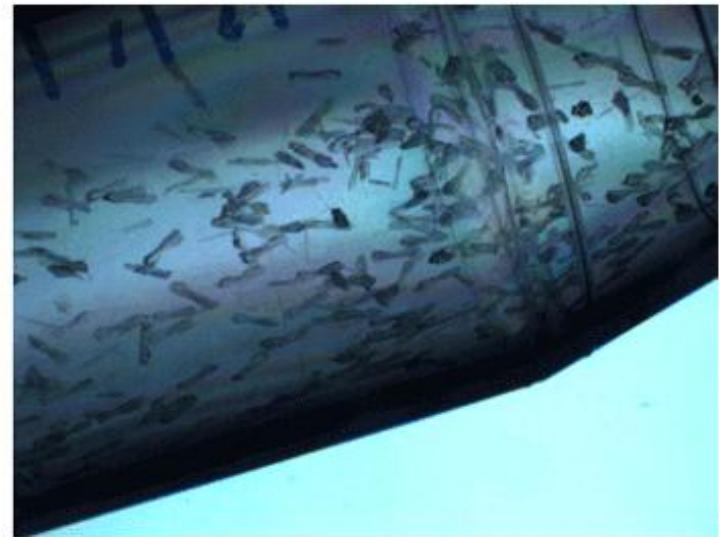
- CCP4, <http://www ccp4.ac.uk/>
- CCP14, <http://www ccp14.ac.uk/>

## Traditional Methods in protein crystallography

- Molecular Replacement
- Isomorphous Replacement

# Second SH3 domain of Ponsin: SH3.2

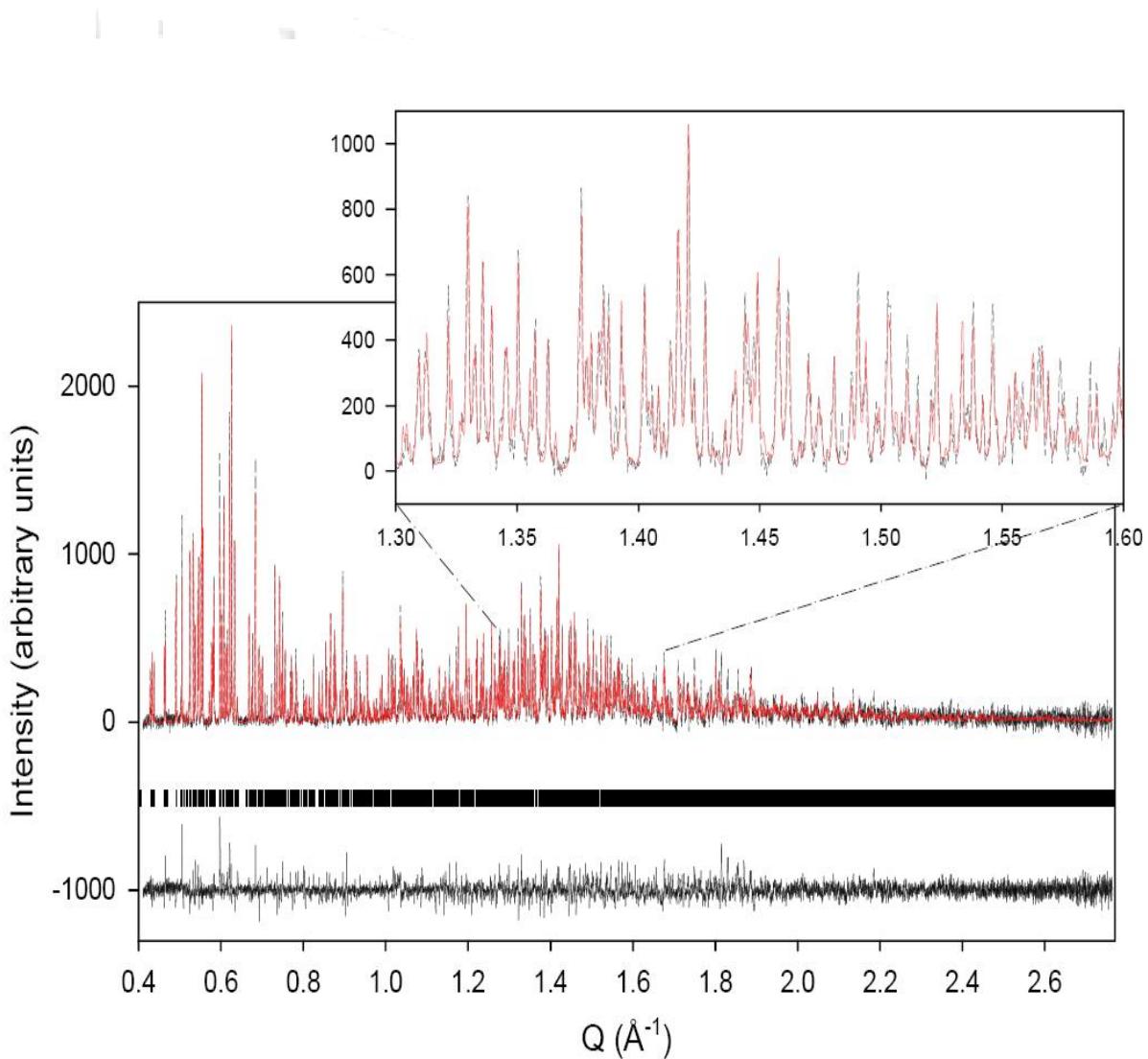
After purification the SH3.2 domain spontaneously formed a microcrystalline material suitable only for powder diffraction measurements



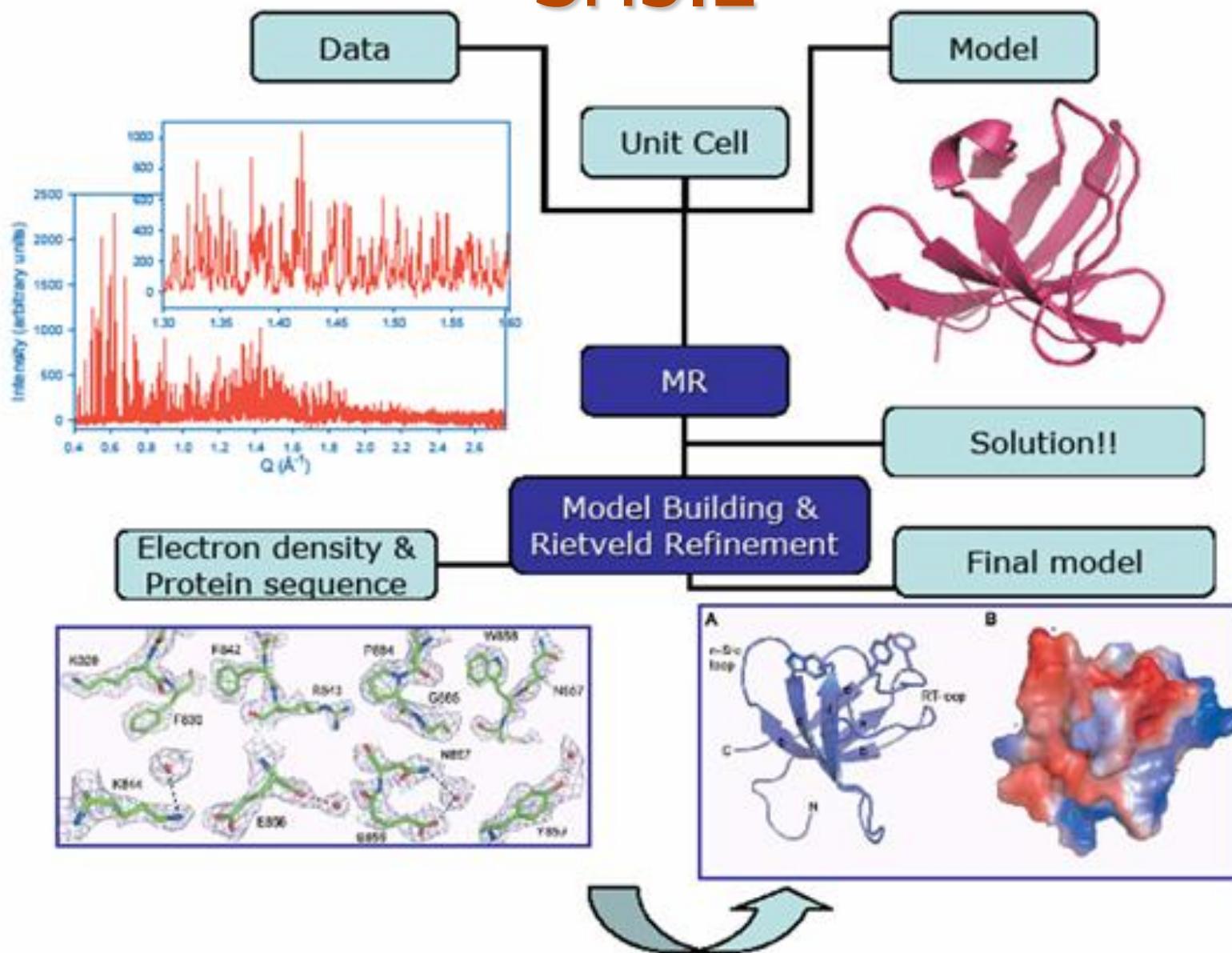
# Data collection



- Two samples A & B
- Spinning glass capillaries  
 $d=1.5\text{ mm}$
- Mounted on the axis of the diffractometer
- Patterns were measured with a period of 2.0 min
- Beam size:  $1.5\text{ mm}^2$   
( $1.5\text{ mm} \times 1.0\text{ mm}$ )
- Photon flux on sample  
 $\sim 3 \times 10^{12}\text{ photons} \cdot \text{s}^{-1} \text{ mm}^{-2}$



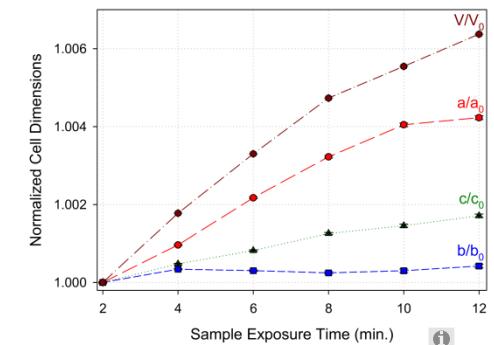
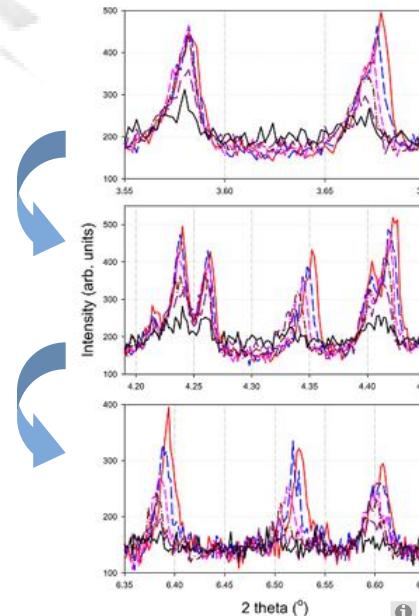
# Procedure followed for data analysis for SH3.2



# Intensity Extraction for SH3.2 domain of ponsin

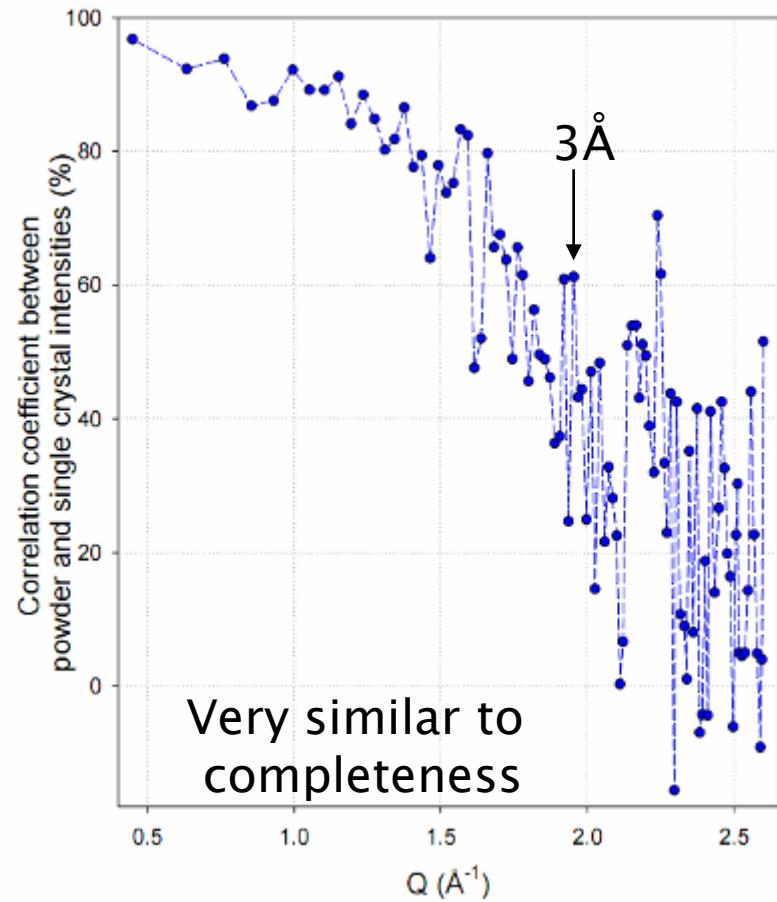
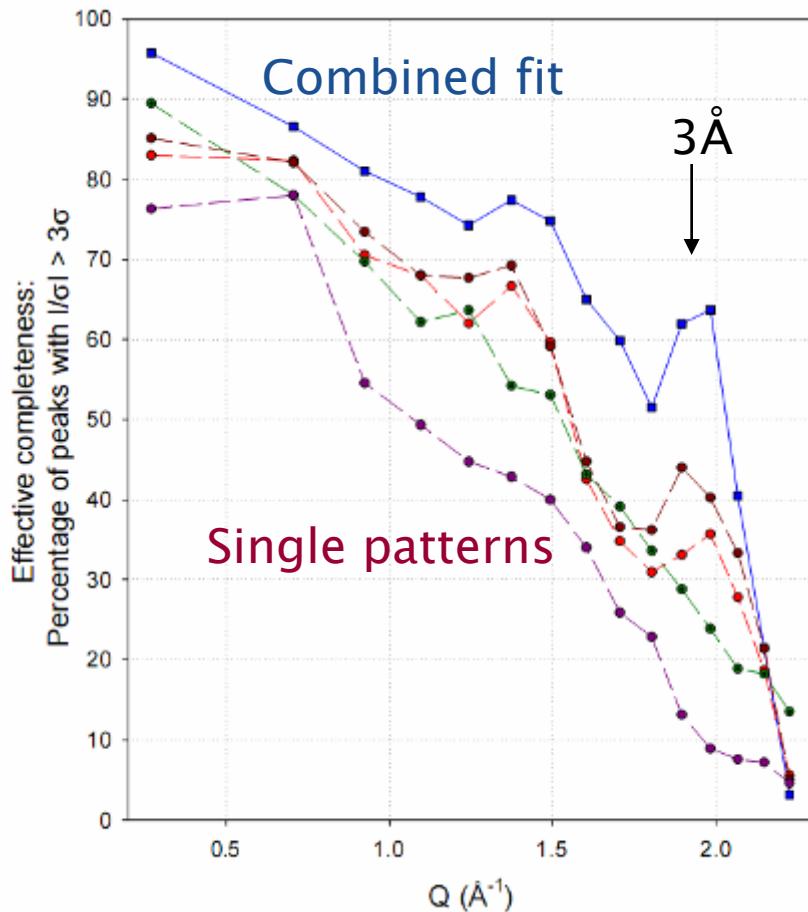
- Anisotropic peak shifts with radiation damage
- Classification of scans using cluster analysis (Pycluster)
- Multi-pattern fit (prodd)
- Use of likelihood function to equi-partition intensities

P<sub>2</sub>12<sub>1</sub>2<sub>1</sub>, a = 24.80, b = 36.38, c = 72.25 Å



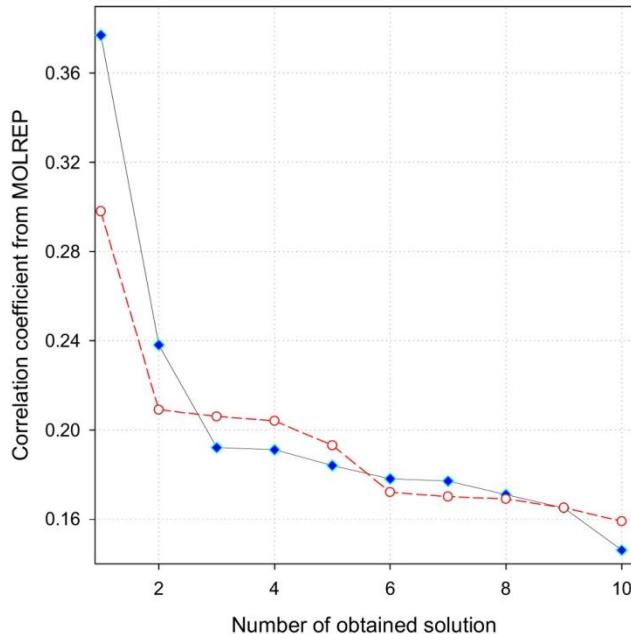
# Quality of extracted intensities

## Completeness & correlations to SX



# Molecular replacement

Solution Score for two different models



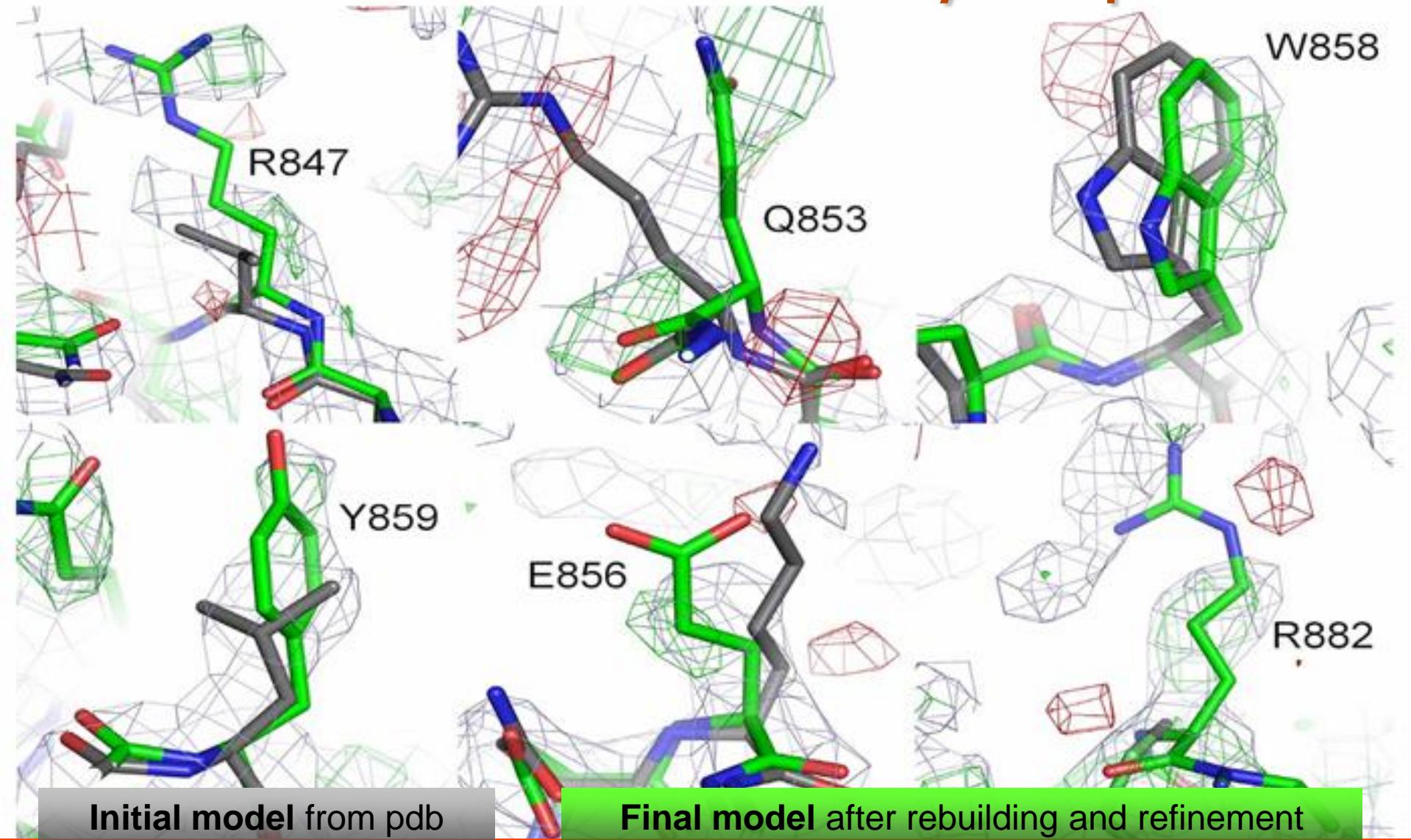
Two models were tested  
Solution always obvious

## Molrep results:

--- Summary ---

S_	RF	TF	theta	phi	chi	tx	ty	tz	TFcnt	Rfac	Scor
S_1_1	1	58.07	144.99	102.00	0.161	0.266	0.181	3.47	0.527	0.377	
S_6_15	2	145.90	144.44	92.50	0.361	0.166	0.366	5.14	0.595	0.238	
S_2_10	3	125.77	-81.54	96.49	0.212	0.061	0.165	2.29	0.627	0.192	
S_4_7	4	51.52	58.03	125.28	0.456	0.267	0.351	3.25	0.612	0.191	
S_7_5	5	132.09	-179.53	72.05	0.167	0.366	0.214	1.40	0.621	0.184	
S_8_10	6	175.99	-179.51	84.91	0.406	0.101	0.391	2.09	0.615	0.178	
S_10_13	7	71.94	-151.69	19.36	0.378	0.216	0.318	1.78	0.625	0.177	
S_5_5	8	62.17	-139.62	106.19	0.279	0.485	0.198	2.21	0.624	0.171	
S_9_11	9	80.99	-179.11	79.75	0.284	0.452	0.132	1.62	0.616	0.165	
S_3_13	10	30.37	76.91	178.57	0.346	0.253	0.137	2.17	0.627	0.146	

# Initial electron density maps

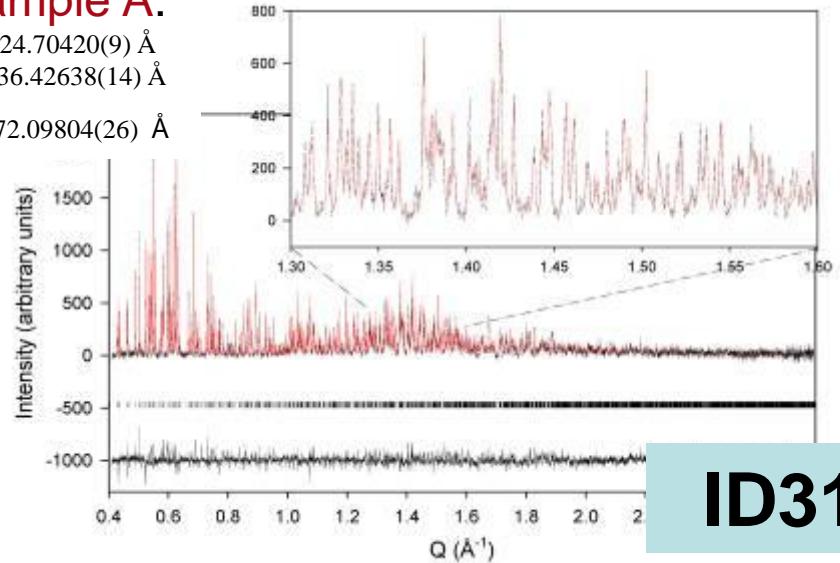


2Fo-1Fc (blue at  $1\sigma$ ) and 1Fo-Fc (red at  $-2.5\sigma$  and green at  $2.5\sigma$ ) electron density maps, as determined directly after the molecular replacement. The residues represented in grey stick carbon atoms correspond to the molecular replacement model used for the calculation of the maps, while the residues in green color carbon atom sticks represent the final refined model.

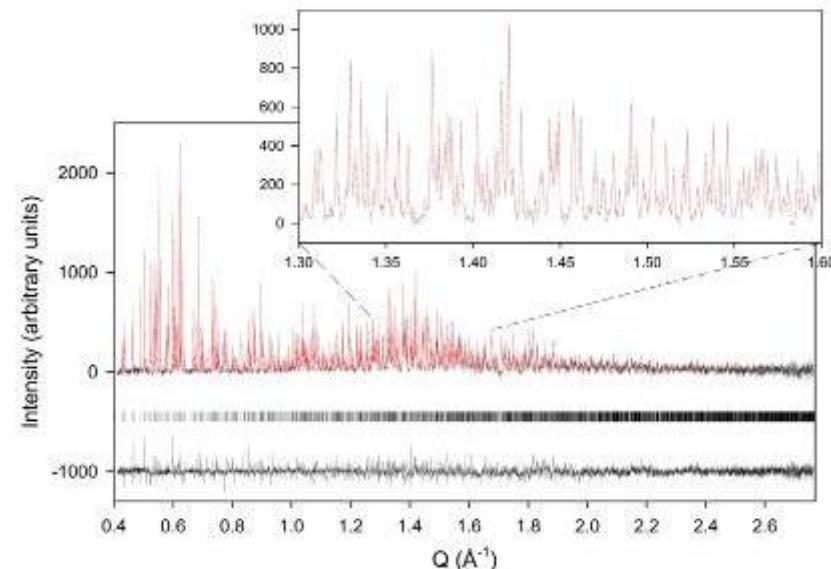
# 4 dataset restrained refinement

Sample A:

$\lambda = 1.252481(32)$  Å, exposure time: 2 min.

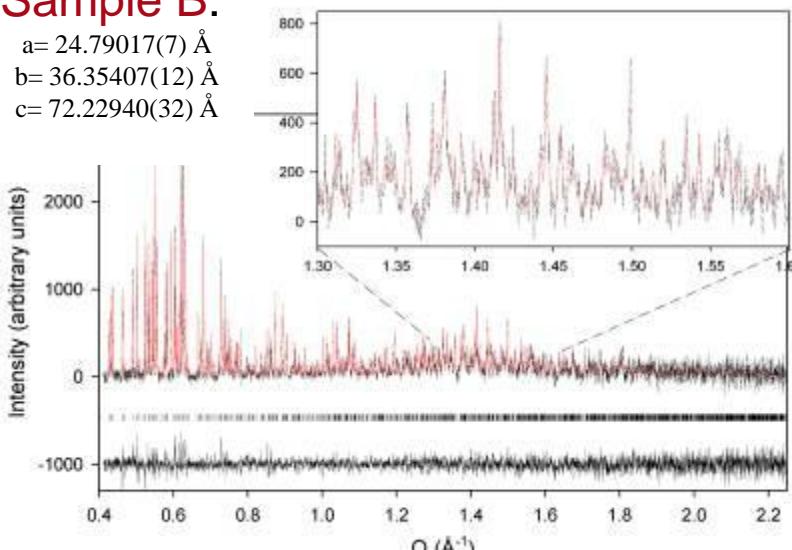


$\lambda = 1.252481(32)$  Å, exposure time: 4 min.

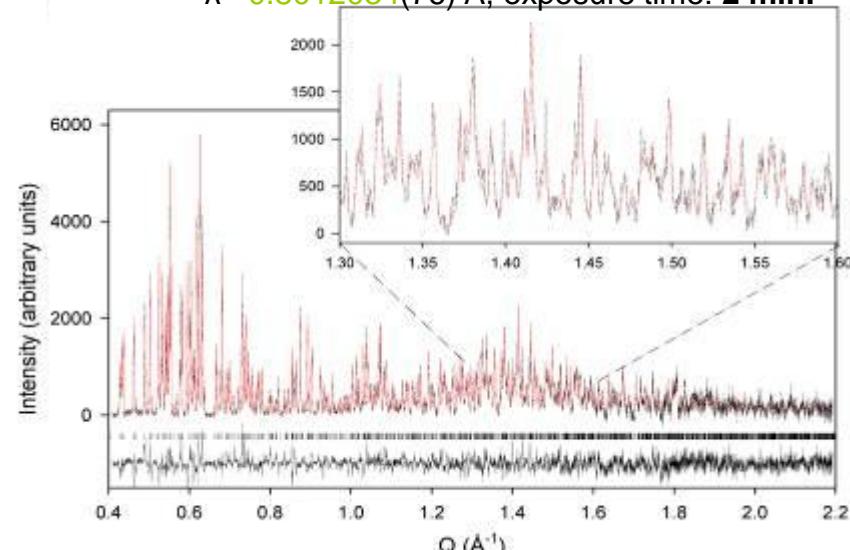


Sample B:

$\lambda = 1.251209(40)$  Å, exposure time: 2 min.



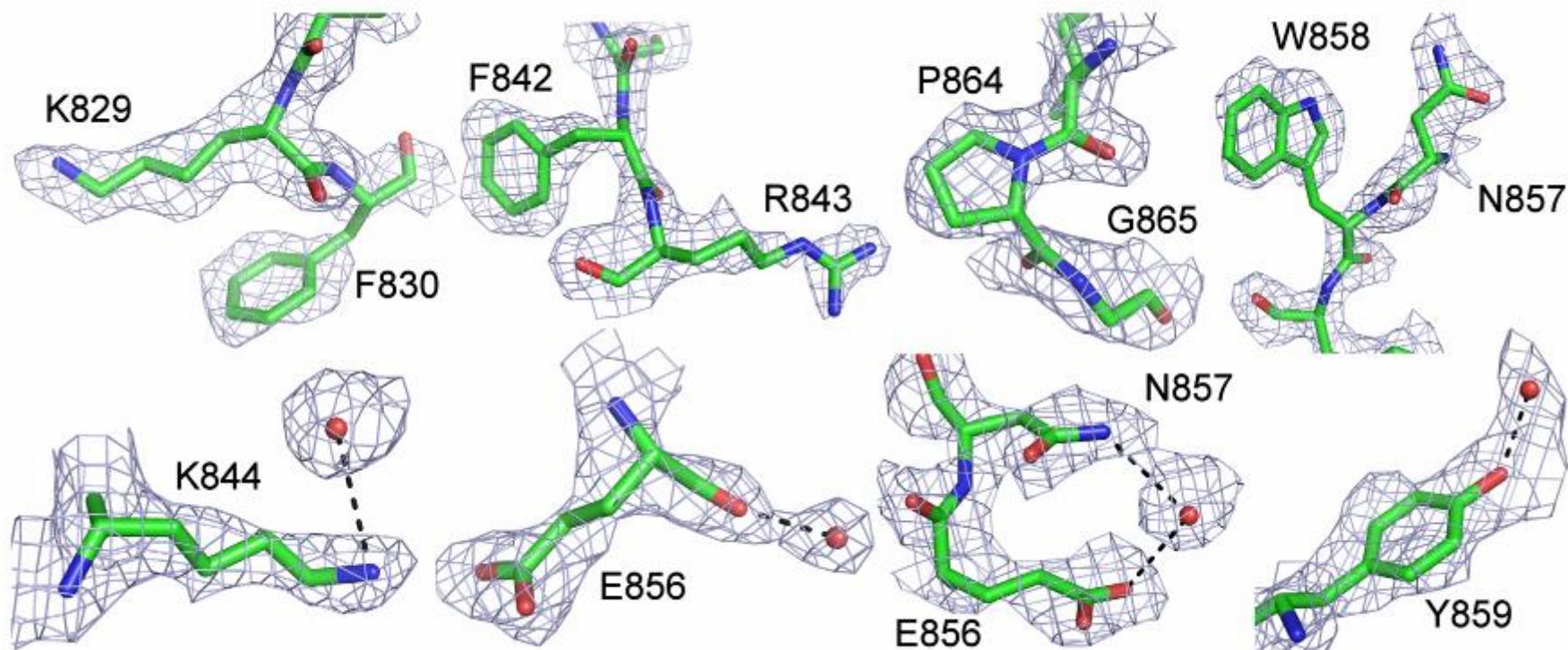
$\lambda = 0.8012034(76)$  Å, exposure time: 2 min.



# The Second SH3 domain of Ponsin

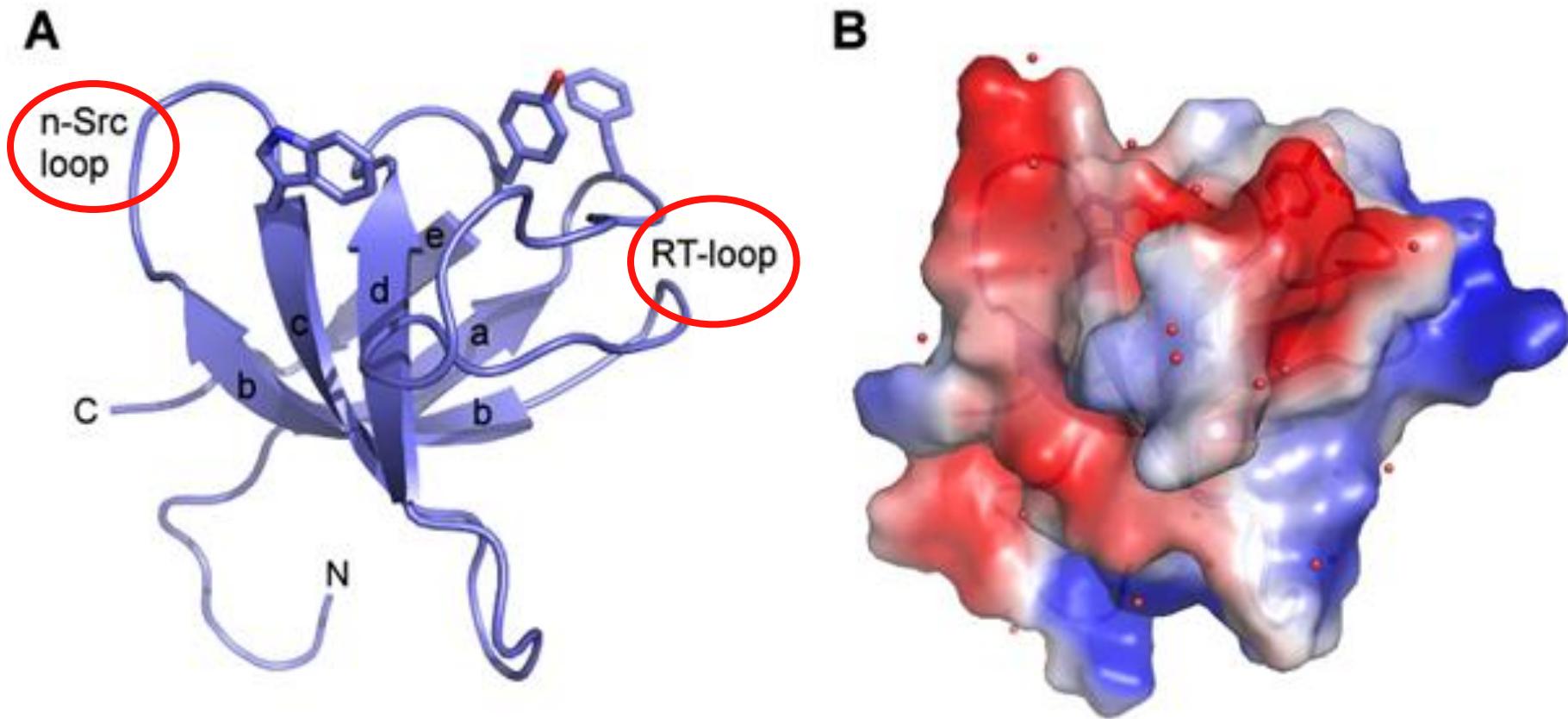
Selected regions of the final refined structural model in stick representation and the corresponding total omit map contoured at  $1\sigma$ .

544 protein atoms and 36 water molecules were identified



# SH3.2: The final model

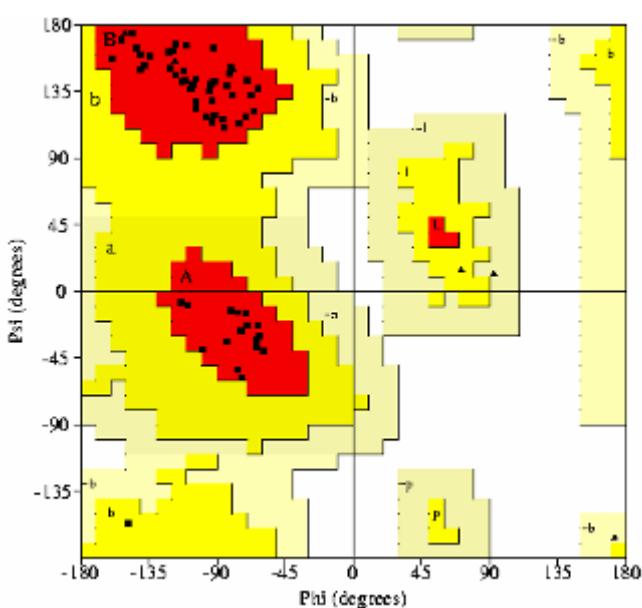
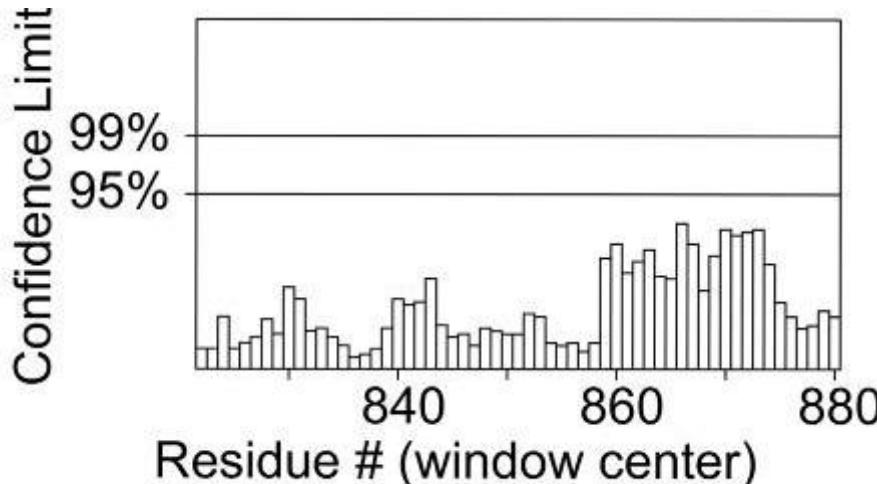
Powder-diffraction structure of the ponsin SH3.2 domain. (A) Ribbon representation of the SH3.2 indicating the secondary structure elements of the domain. The main hydrophobic residues of the binding interface as well as the positions of the n-Src and RT loops are indicated. (B) Electrostatic potential representation of domain identifying additionally the water molecules as red spheres.



# Structure Validation

## ERRAT:

<http://nihserver.mbi.ucla.edu/ERRATv2/>



## PROCHECK:

<http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.htm>

+-----<<< P R O C H E C K S U M M A R Y >>>-----+

| final2 2.8 67 residues |

| Ramachandran plot: 98.2% core 1.8% allow .0% gener .0% disall |

| Gly & Pro Ramach: 0 labelled residues (out of 10) |

| Chi1-chi2 plots: 0 labelled residues (out of 43) |

| Main-chain params: 6 better 0 inside 0 worse |

| Side-chain params: 5 better 0 inside 0 worse |

| Residue properties: Max.deviation: 2.7 Bad contacts: 0 |

+| Bond len/angle: 4.8 Morris et al class: 1 1 3 |

| G-factors Dihedrals: -.24 Covalent: -.17 Overall: -.18 |

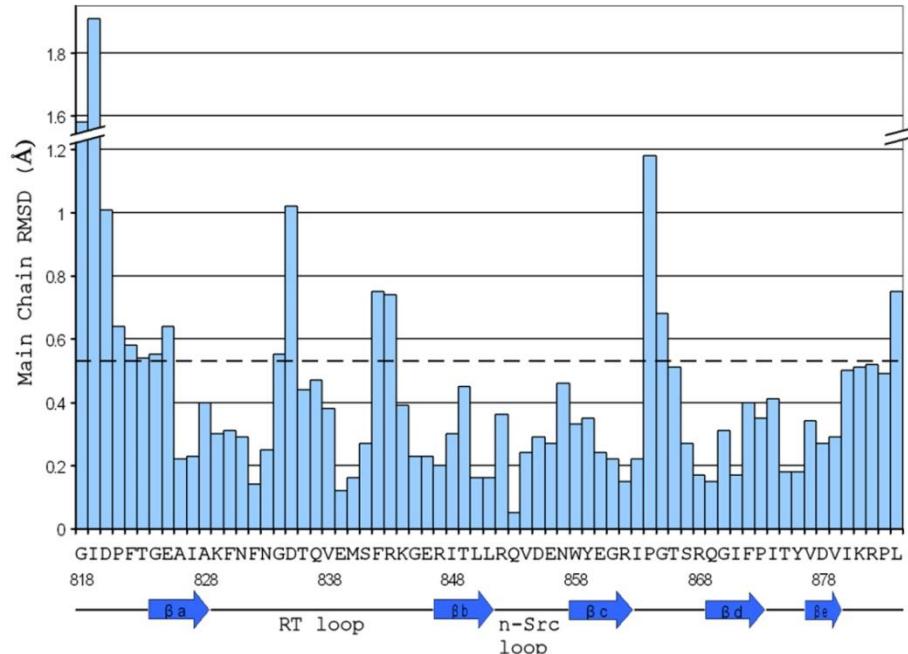
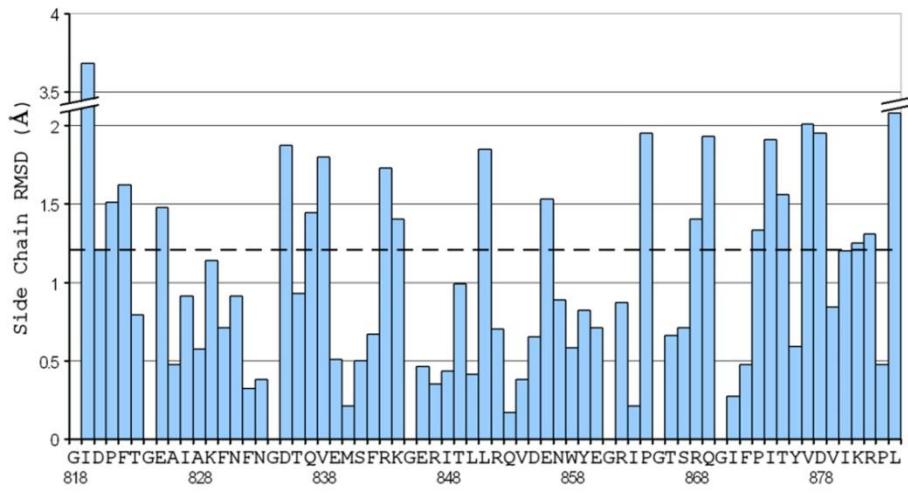
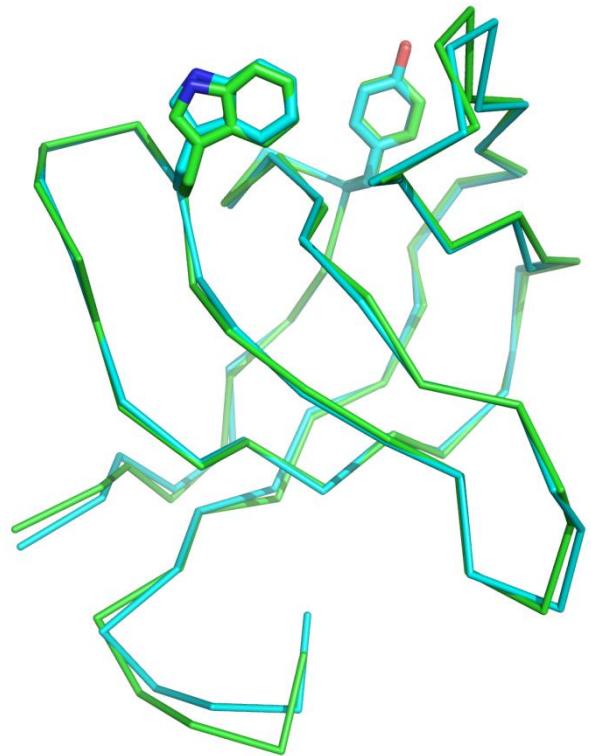
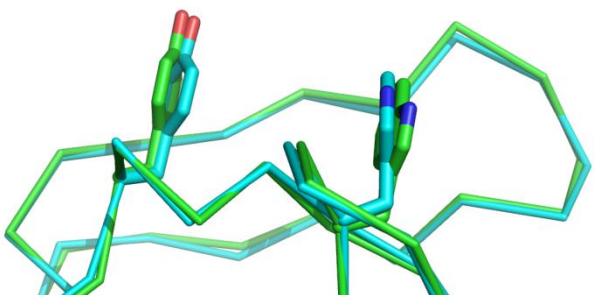
| M/c bond lengths: 100.0% within limits .0% highlighted |

| M/c bond angles: 80.1% within limits 19.9% highlighted |

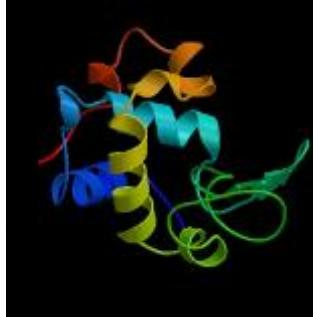
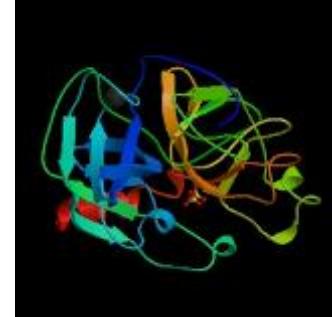
+| Planar groups: 95.5% within limits 4.5% highlighted |

+ May be worth investigating further. \* Worth investigating further.

# Comparison with the single crystal model

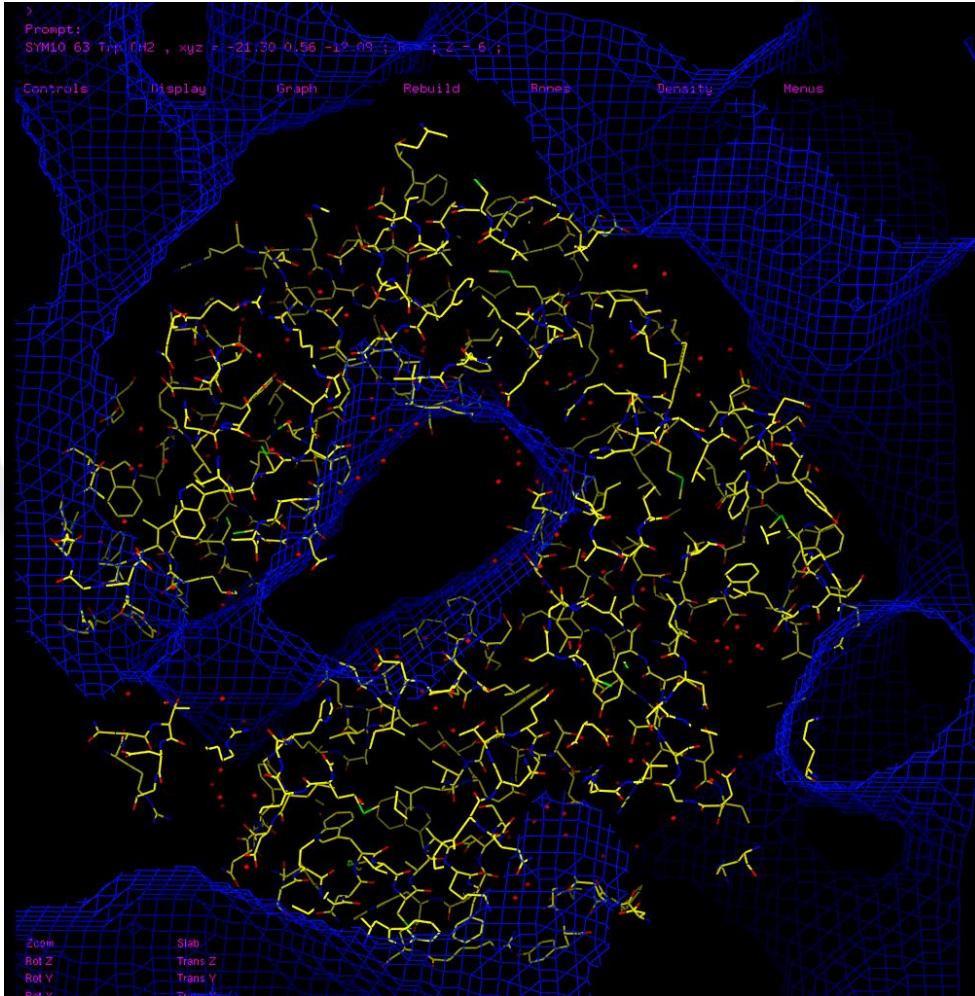


# Test systems

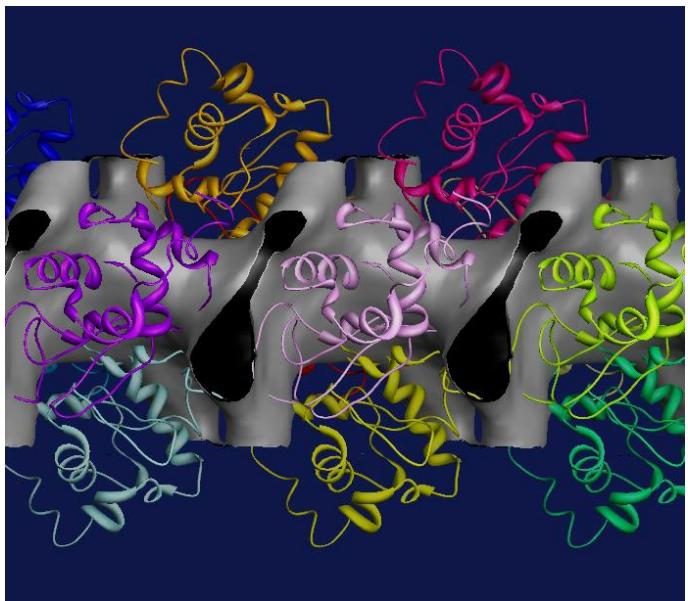
protein	<p><b>Hen egg-white Lysozyme (HEWL)</b></p> 	<p><b>Pancreatic porcine Elastase (PPE)</b></p> 
Molecular weight	<b>14.4 kDa</b>	<b>24.8 kDa</b>
Unit-cell (Å)	<b>a=b=79.2      c=38.0</b>	<b>a=51.8   b=57.9   c=75.3</b>
Space-group	<b>P4<sub>3</sub>2<sub>1</sub>2</b>	<b>P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub></b>
Relative root mean square intensity change  (Crick and Magdoff)	<b>Gd (Z=64) : 0.40</b>	<b>U(Z=92) : 0.43</b>

# Low resolution phasing in Gd derivative of lysozyme

**METHOD: Single Isomorphous Replacement (SIR)**

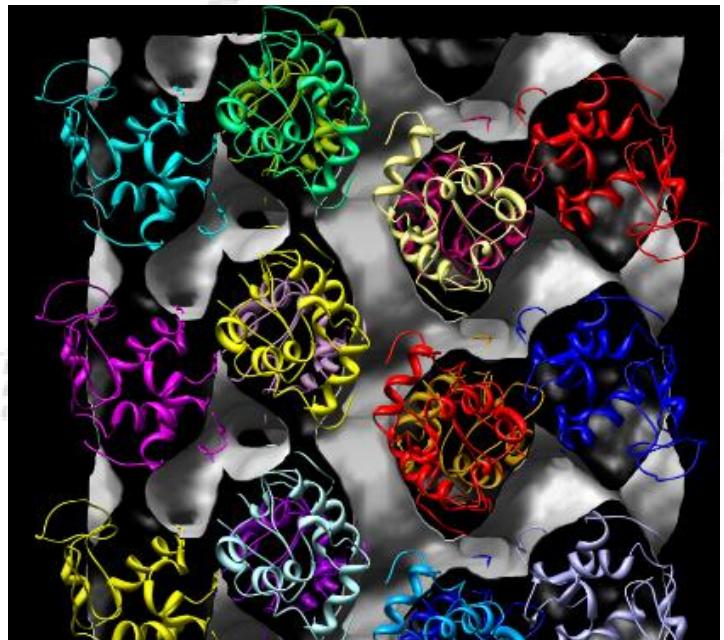


- Sites from shelxd
- Refined in sharp
- Followed by density modification
- Figure shows solvent mask from density modification

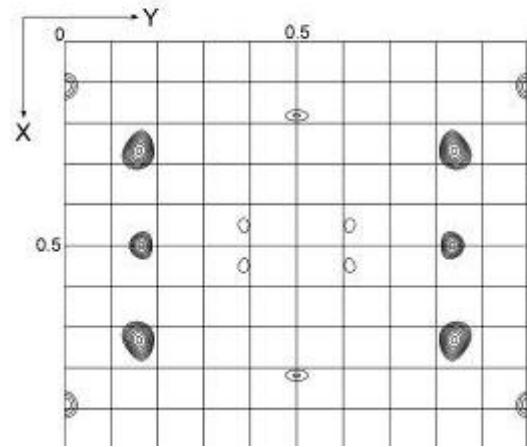
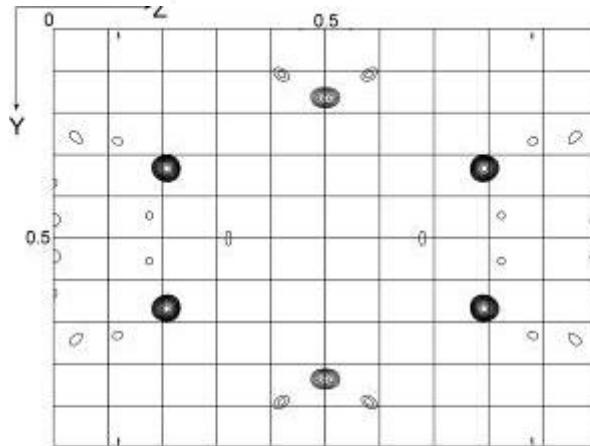


Solvent channel found from map

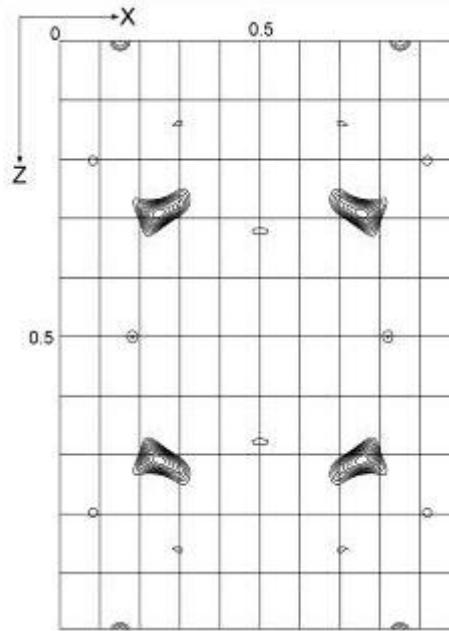
Gd-HEWL



# Heavy atom detection



Harker sections in  
the Patterson map  
for elastase.



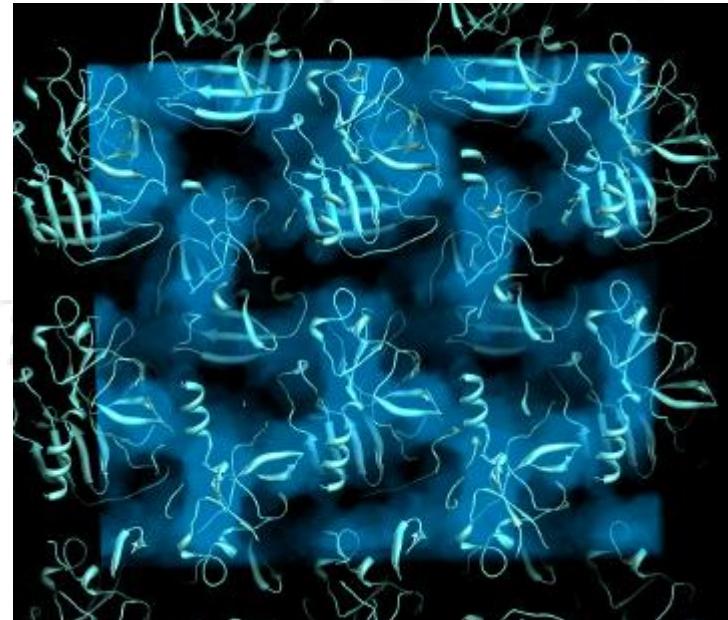
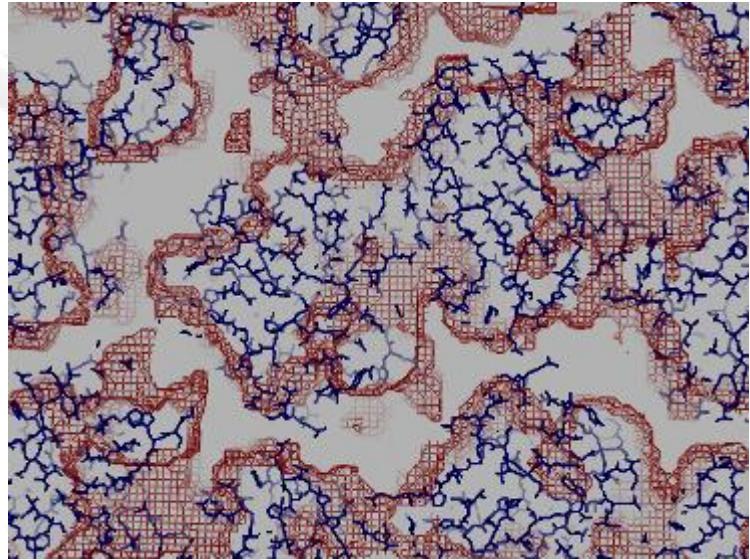
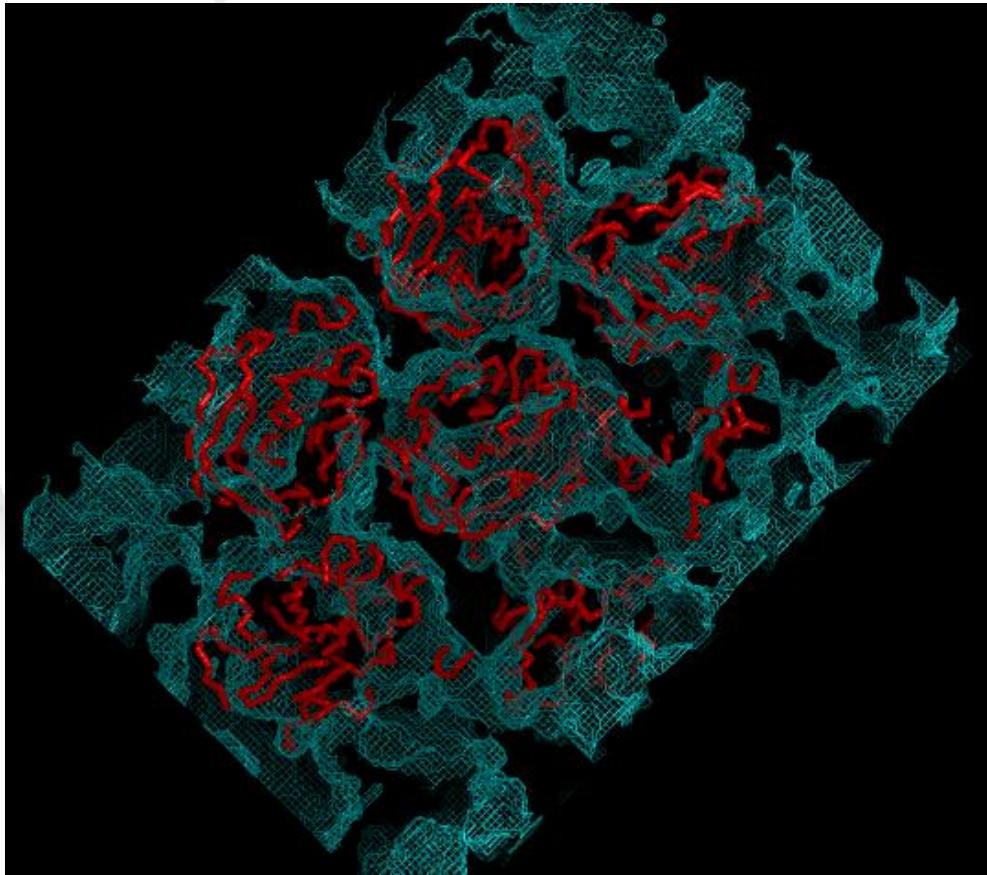
Elastase

Isomorphous  
Patterson map

Lysozyme

Direct methods  
(SHELXD)

# Computed maps for Elastase-U



J. P. Wright et al. J. Appl. Cryst. (2008)

Margiolaki, I. & Wright, J. P., Acta Cryst. (2008). A64, 169–180

# The quality of the density maps

- The correlations with the calculated maps are better for PPE than for HEWL.
- Can we improve them by using more derivatives?

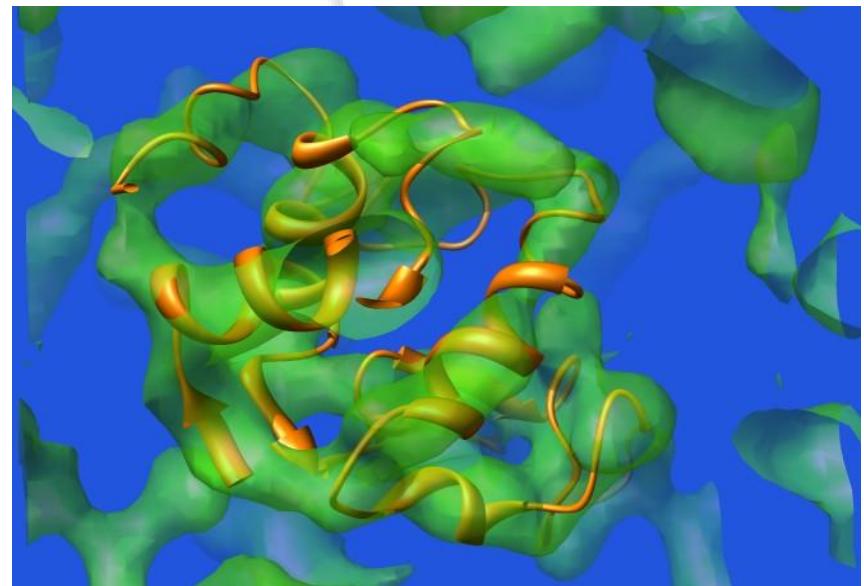
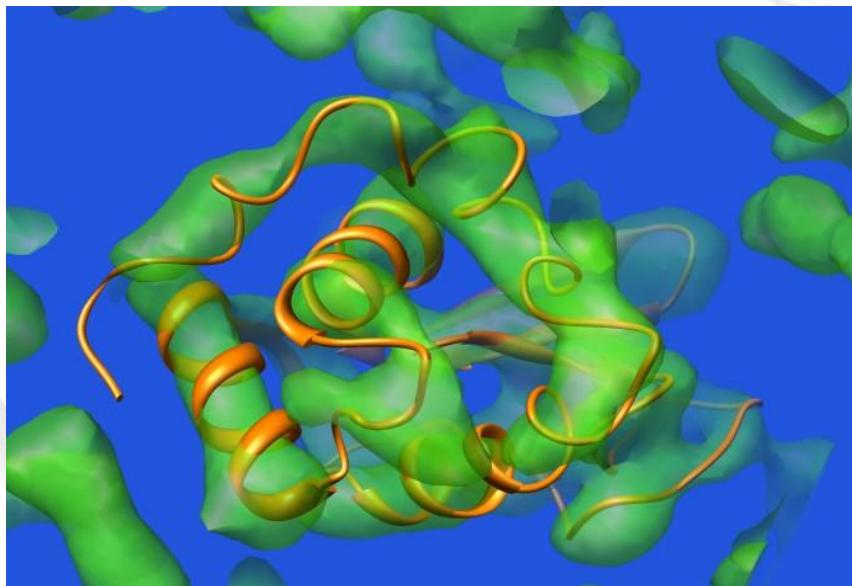
YES

inf	9.85	0.5459	0.4555	0.7686
9.85	7.82	0.3388	0.3128	0.7103
7.82	6.84	0.4594	0.2355	0.4796
6.84	6.21	0.4009	0.4236	0.6045
6.21	5.77	0.3032	0.2317	0.4164
5.77	5.43	0.4432	0.2046	0.4848
5.43	5.16	0.3432	0.0128	0.1325
5.16	4.93	0.2994	-0.0573	0.0597
4.93	4.74	0.3237	0.0508	0.1894
4.74	4.58	0.2373	-0.2324	0.0076
4.58	4.44	0.2382	-0.1204	0.2409
4.44	4.31	0.2573	0.0888	0.0985
4.31	4.2	0.2415	-0.0687	0.0355
4.2	4.09	0.2037	-0.0176	0.2504
4.09	4	0.139	0.2801	0.0489

Resolution (Å)

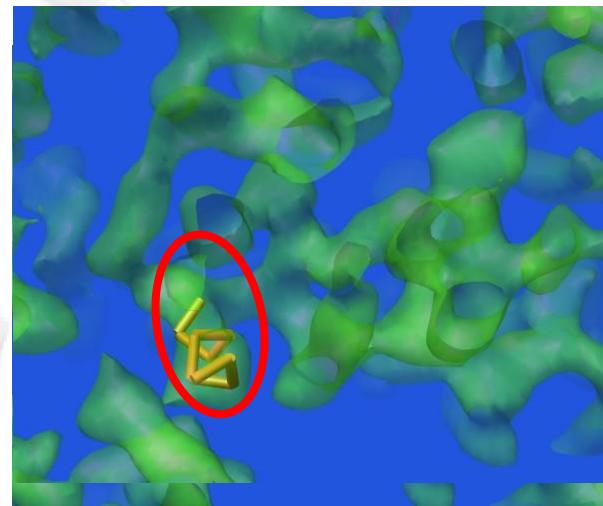
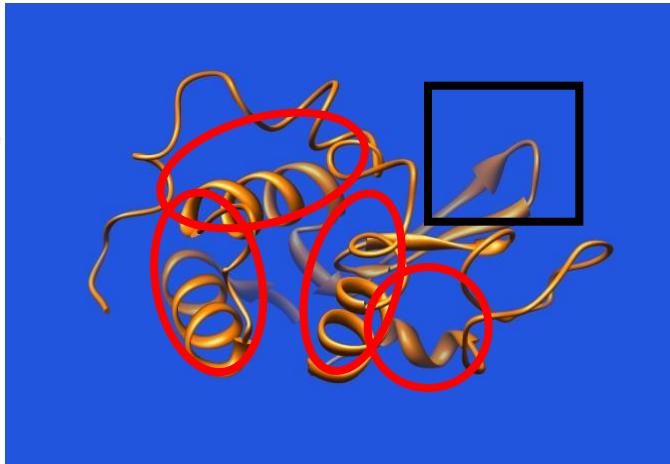
PPE   HEWL   HEWL  
SIR   MIR

# The HEWL MIR density map



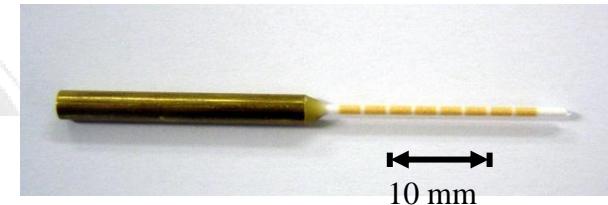
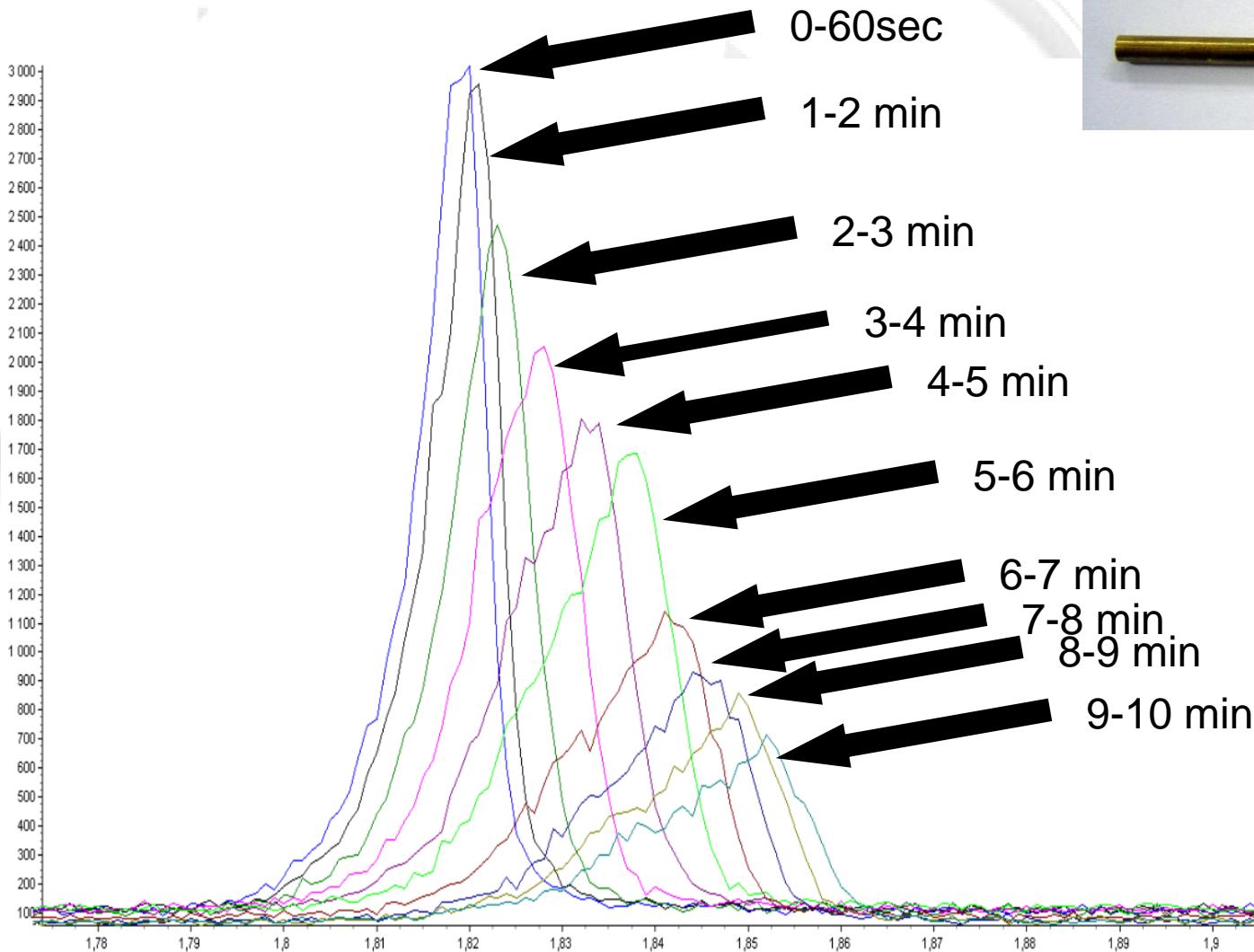
*Images created by Sebastian Basso*

# The secondary structure



- Using FFFEAR we are able to locate all the four helices of the lyzosome structure out of the five best fitted helices found by the program using the standard fragment library.

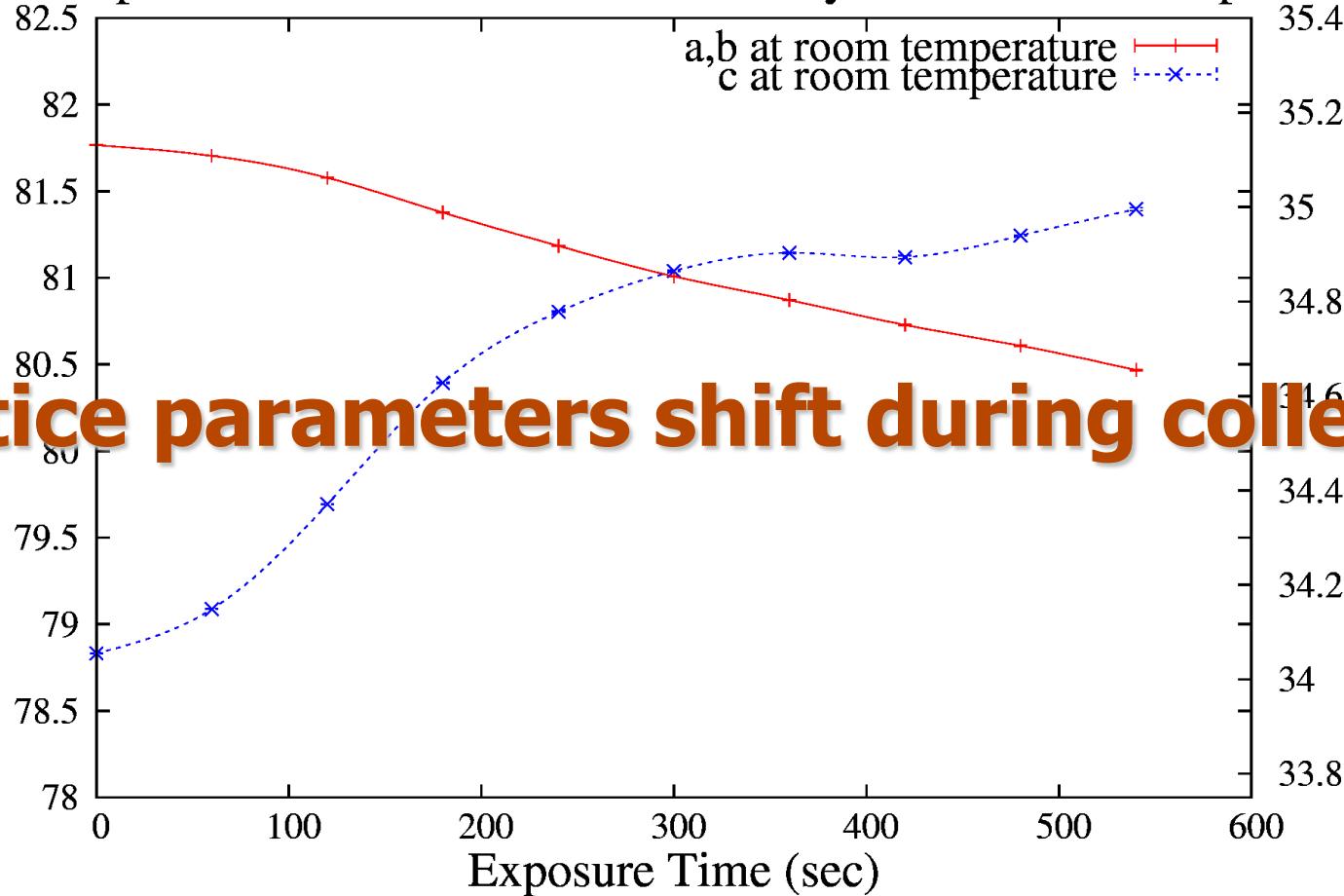
# Effect of radiation damage at Room temperature



Courtesy: Yves Watier

# Effect of radiation damage at Room temperature

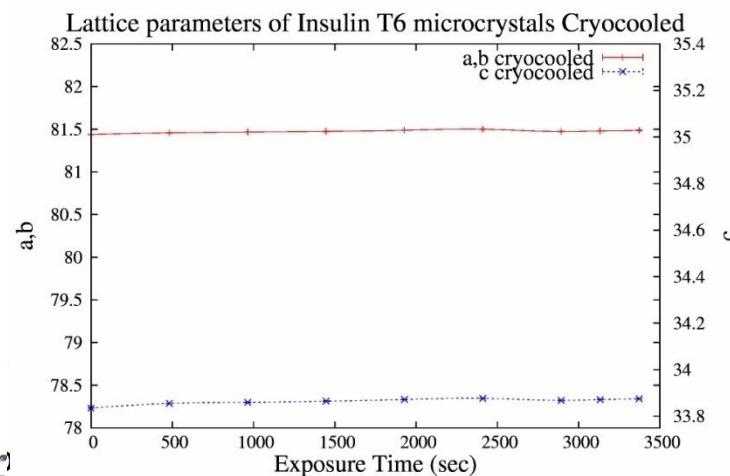
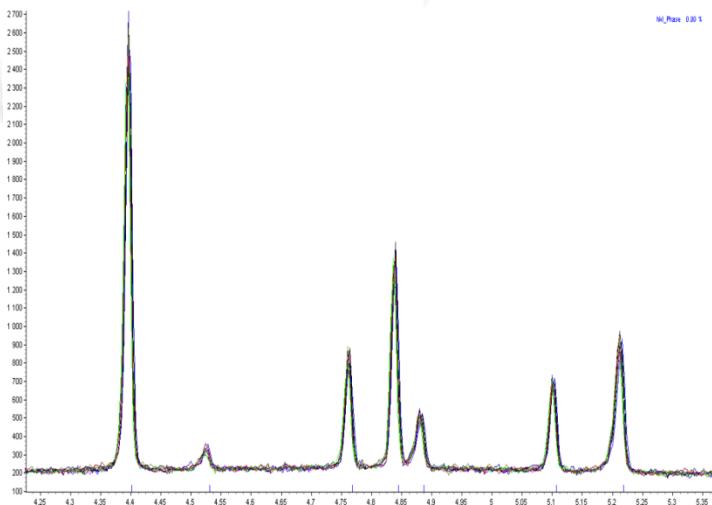
Lattice parameters of Insulin T6 microcrystals at room temperature



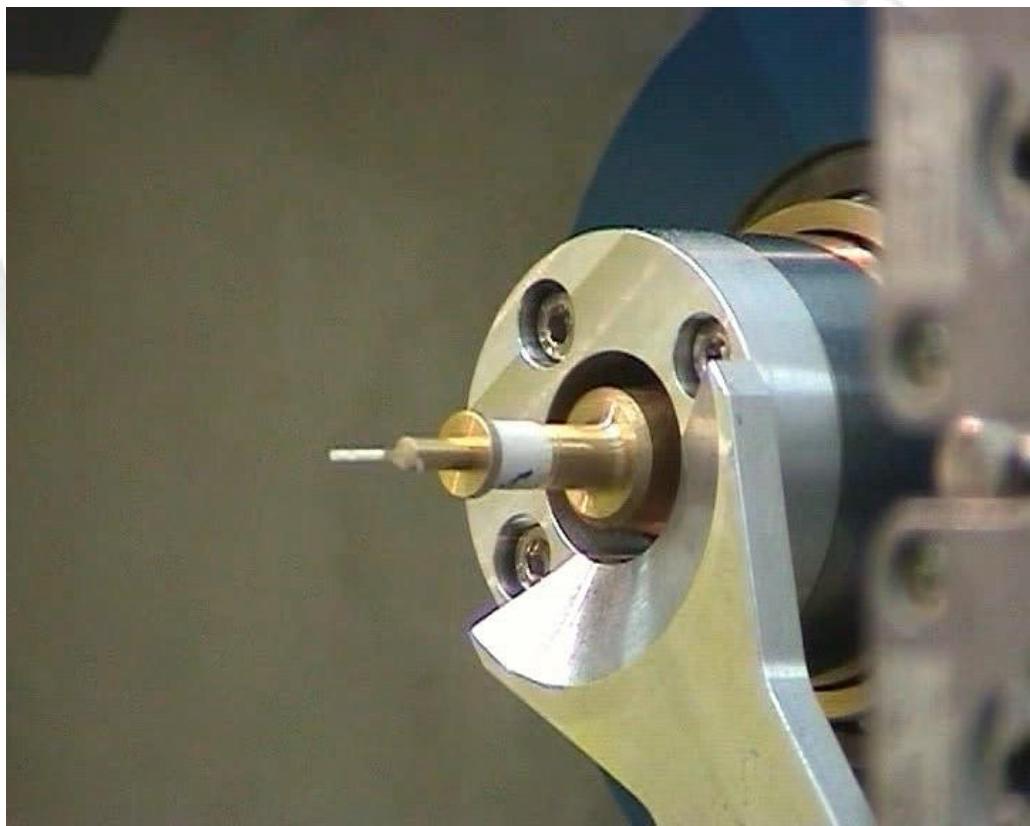
**Lattice parameters shift during collection**

# Cryocooling- Goals

- D-spacing useful is larger.
- No ice peaks.
- Longer collection is possible.
- No radiation damage effects on lattice parameters while cooled.
- Just one position in the capillaries is shot, reducing greatly the amount of sample needed.



# Cryocooling

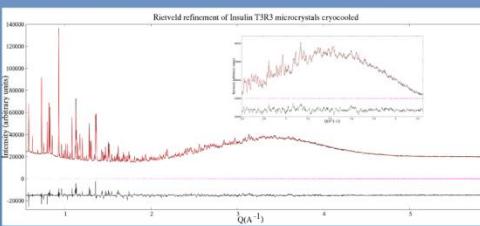
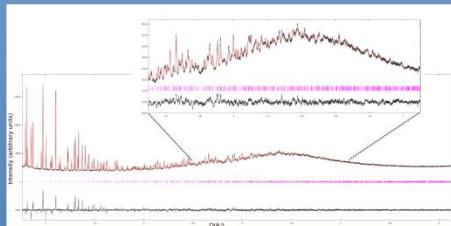


- Vast majority of single crystal structures collected at low temperature
- Radiation damage reduced
- Small reduction of B-factors

Translating sample spinner at ID31 (Fitch/Rossat)

# Successful cryocooling of Insulin

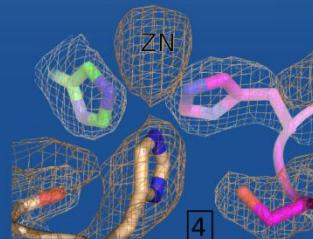
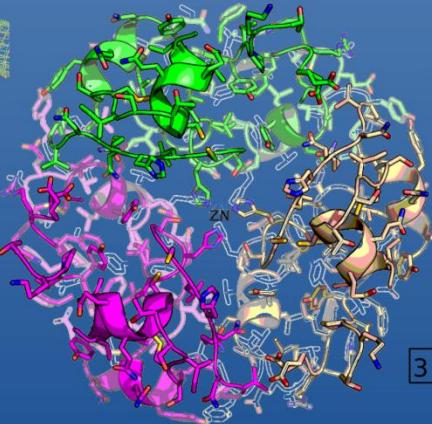
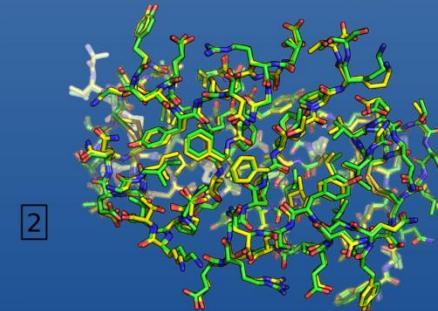
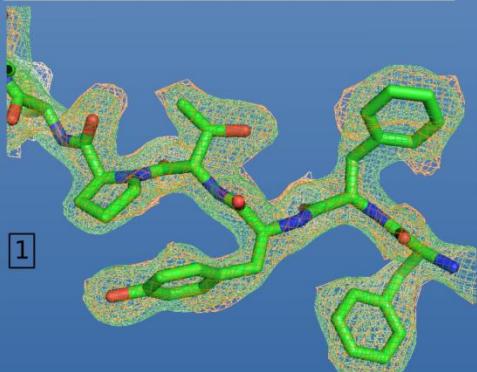
Rietveld refinement for Insulin T6 and T3R3 cryocooled at 100K.



Space group and lattice parameters has been determined with Topas 4.

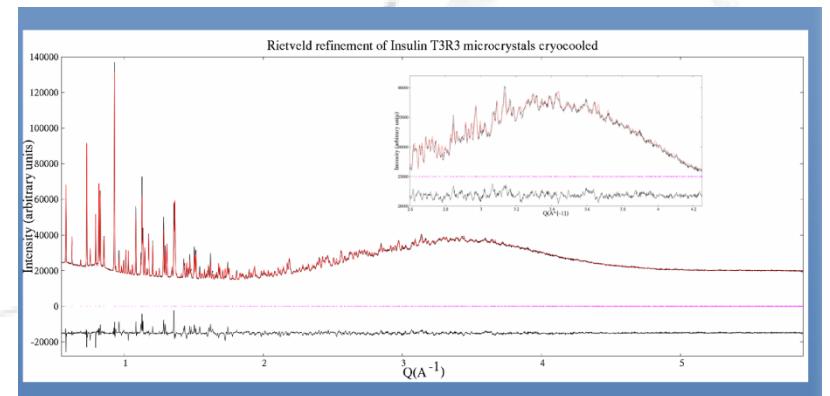
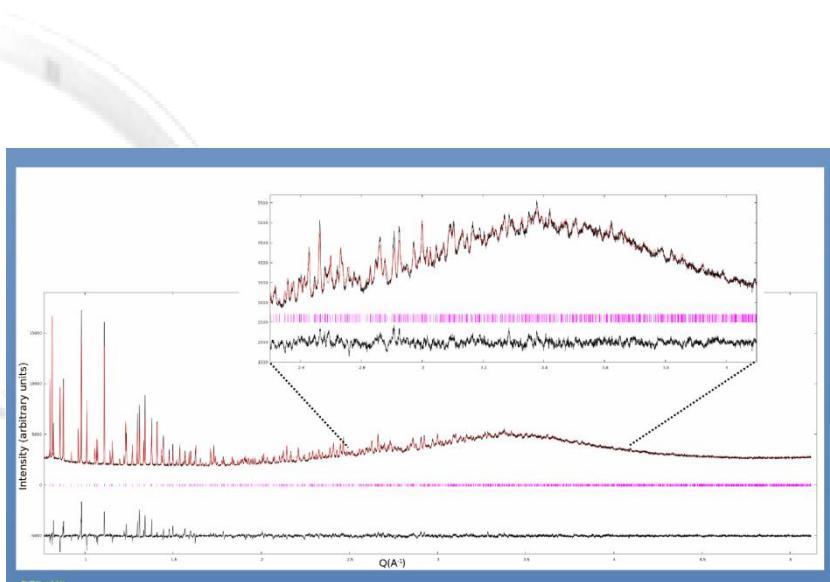
The observed intensities were extracted using prodd [2], in order to perform a molecular replacement with molrep[3].

The Rietveld refinement (atomic positions) has been done with the software GSAS[4][5], then fine tuned with coot using total omit maps computed with sfcheck[6].



[1] Detail of Human Insulin T6 model.  
[2] Difference between original model (1MSO) and refined model of human Insulin T6.  
[3] [4] Coordination of 3 dimers of Insulin around a zinc atom.

[1][3][4] The green maps are fobs contoured at 1 sigma, the orange maps are 7 cycles omit map contoured at 1 sigma.

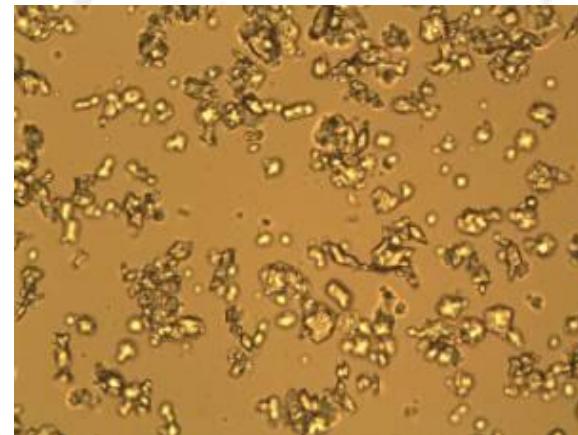


Y. Watier et al.

Acta Cryst. (2008). A64, C312

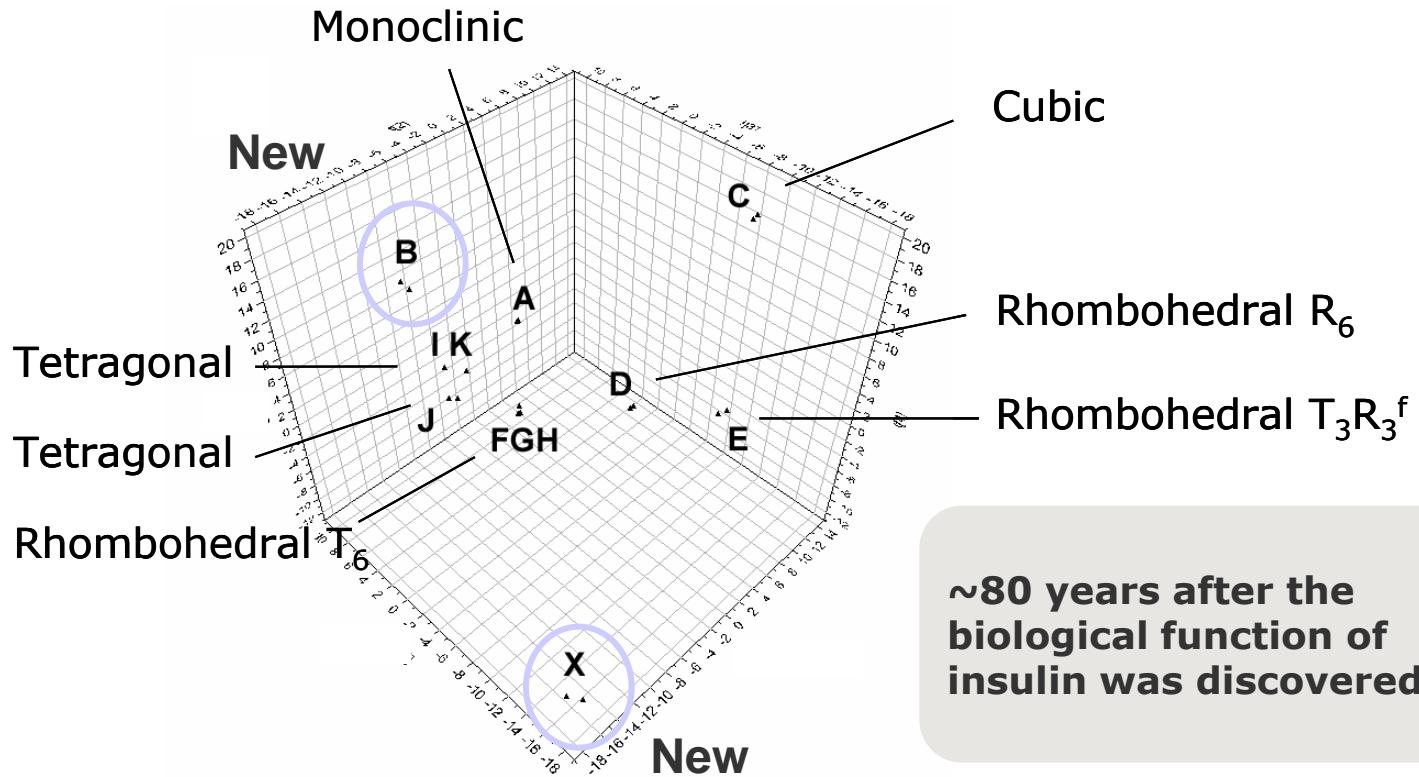
# Industry

- Novo Nordisk – Copenhagen  
(Human Insulin)



# Principal Component Analysis (PCA)

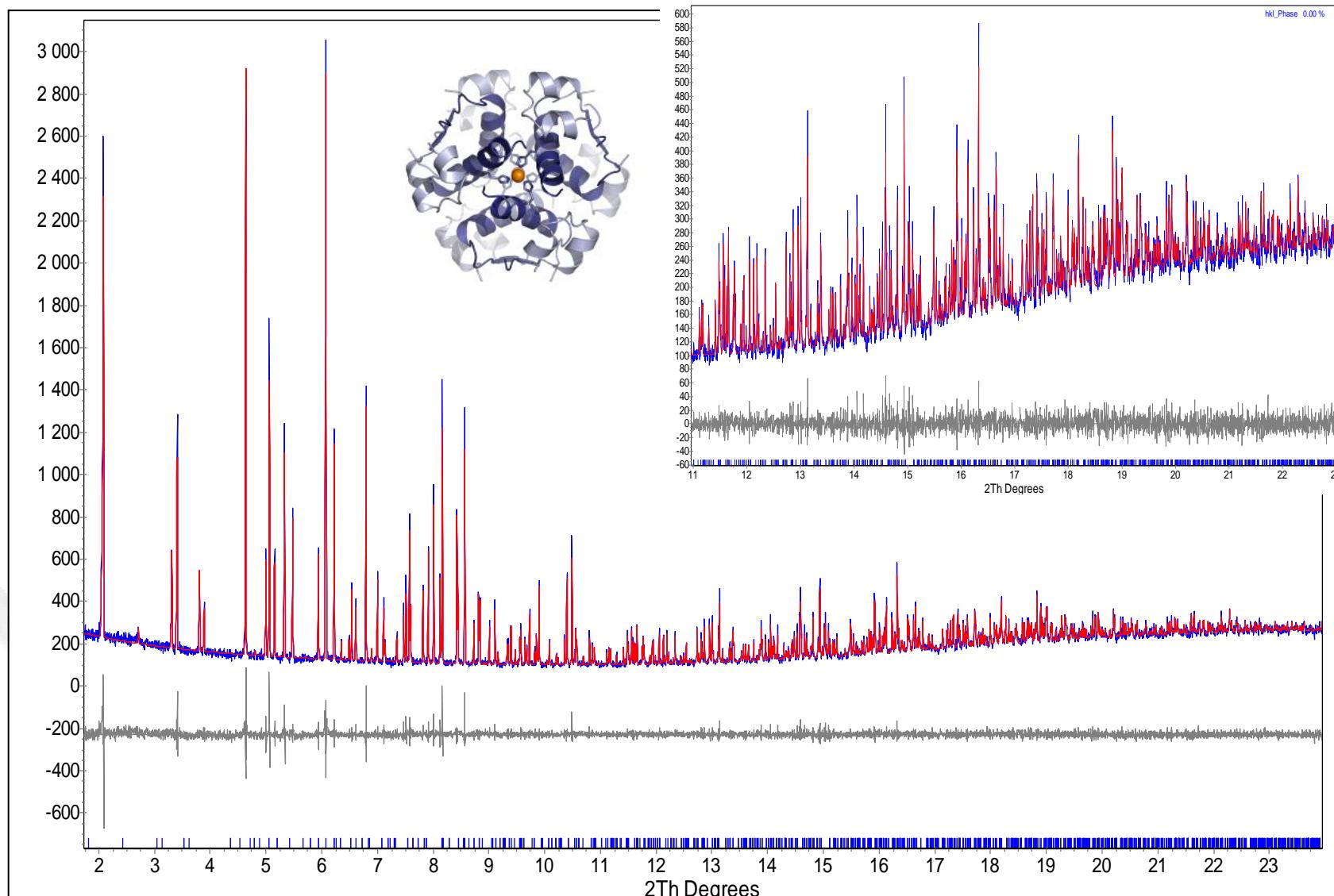
- New insulin crystal forms identified



~80 years after the  
biological function of  
insulin was discovered

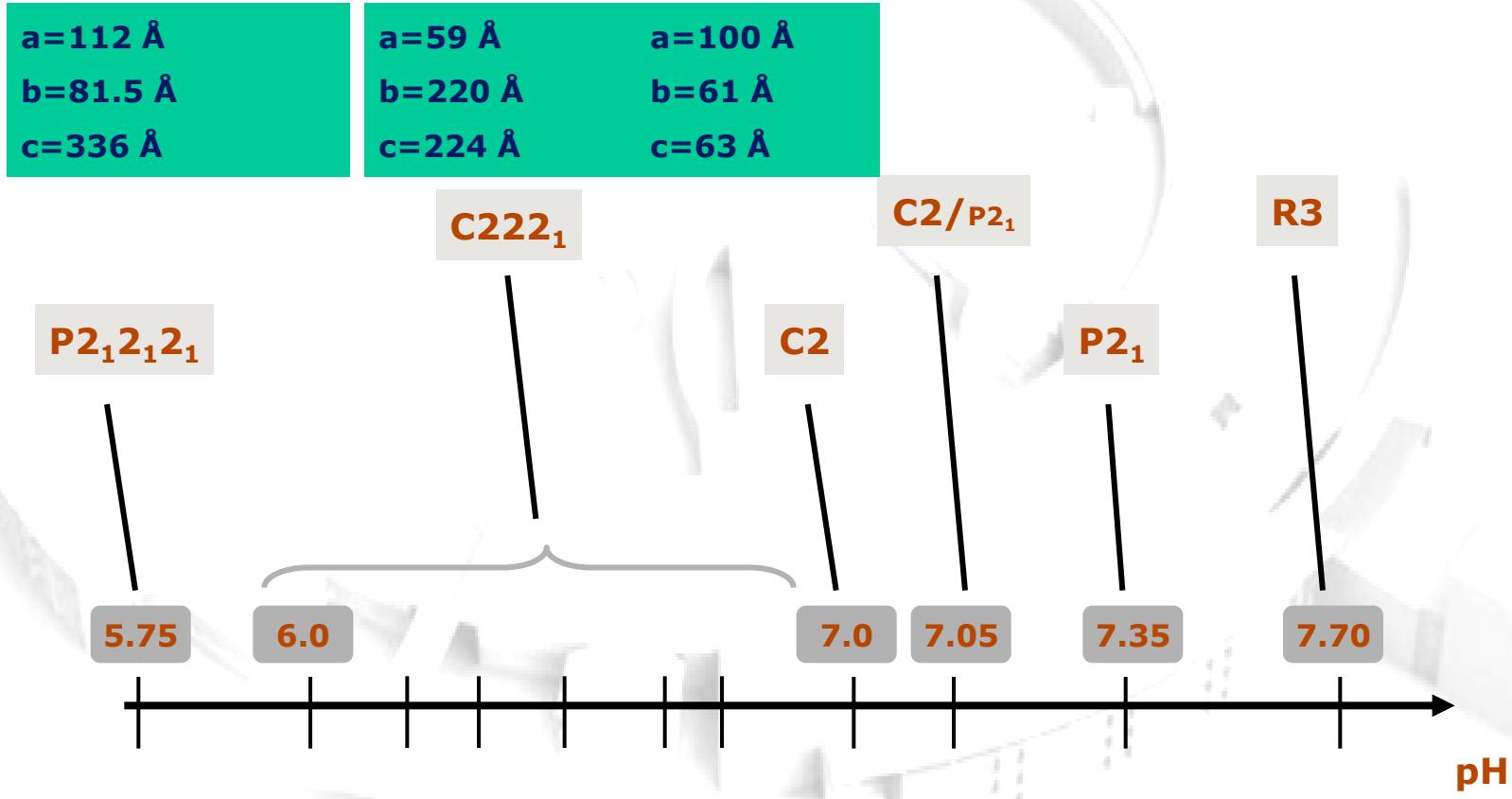
- Norrman, M. et al., (2006) J. Appl. Cryst. 39 391-400
- Margiolaki, I. & Wright, J. P., Acta Cryst. (2008). A64, 169–180
  - Knight, L. et al. (In Preparation)

# T6 Human Insulin



ID31-  $\lambda = 1.24965(2)$  Å  
R3,  $a = 81.7531(4)$  Å,  $c = 34.0688(3)$  Å

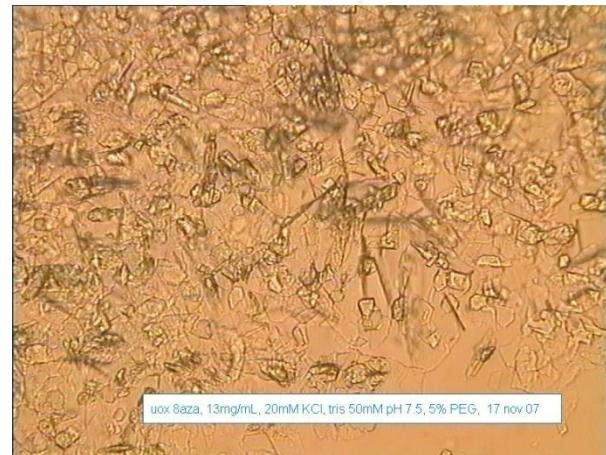
# Phase transitions as a function of pH



- Norrman, M. et al., (2006) J. Appl. Cryst 39 391-400
- Margiolaki, I. & Wright, J. P., Acta Cryst. A (In Press)
- Knight, L. et al. (In Preparation)

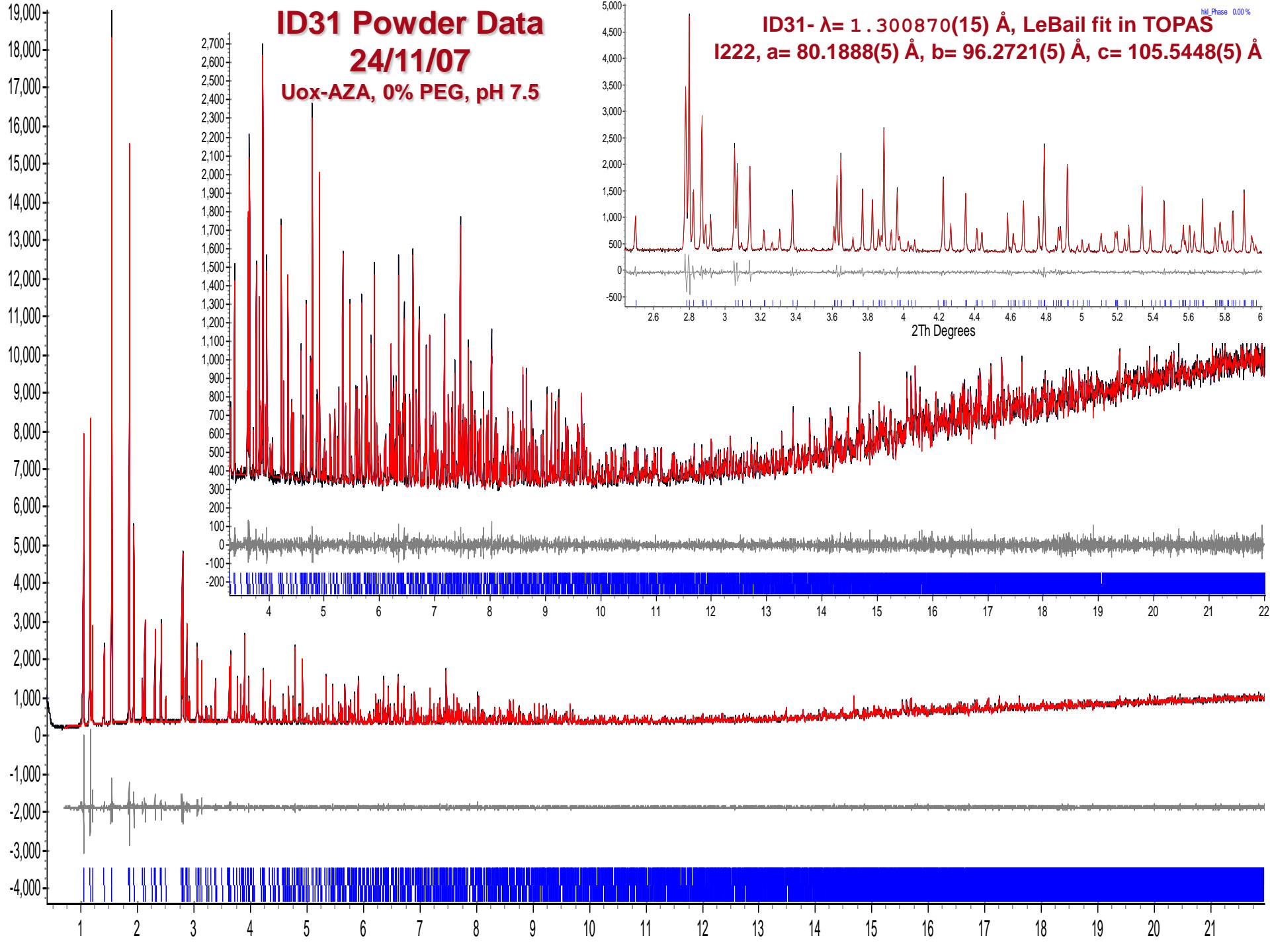
# Industry

- Sanofi Aventis – Paris  
(Urate Oxidase)

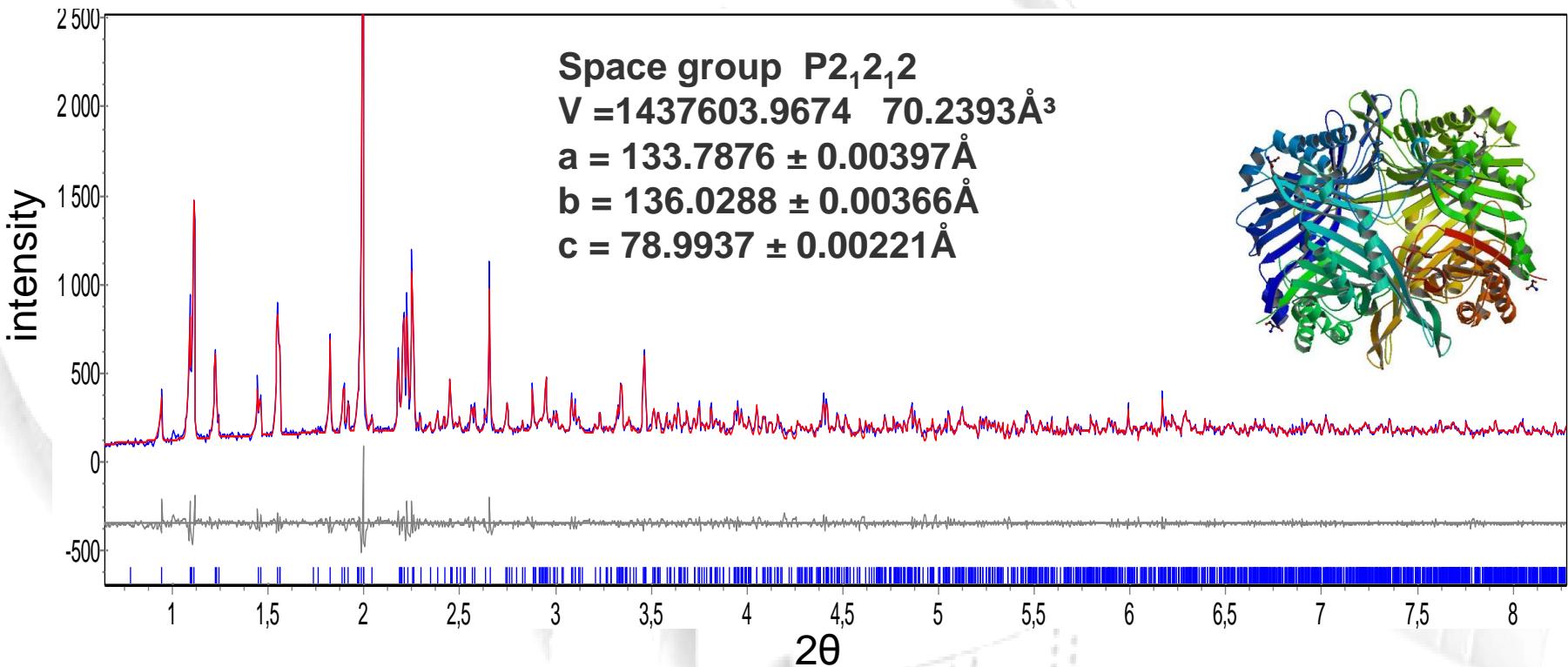


# ID31 Powder Data 24/11/07 Uox-AZA, 0% PEG, pH 7.5

ID31-  $\lambda = 1.300870(15)$  Å, LeBail fit in TOPAS  
I222,  $a = 80.1888(5)$  Å,  $b = 96.2721(5)$  Å,  $c = 105.5448(5)$  Å



# New phase of urate oxidase



# Current Projects

- Characterisation of protein structures for antiviral drug design, AFMB Marseille.  
(Replication proteins from emerging viruses: Chikungunya, SARS, H1N1...)
- Studies of proteins from insect viruses/ virions from the Baculoviridae family, Univ. of Auckland. (CPV's)
- Amyloid- like synthetic peptide analogues of silkworm chorion proteins, Univ. of Athens.

# Acknowledgments

## ESRF

Andy Fitch (ID31)

Jon Wright (ID11)

Yves Watier (ID31)

Sotonye Dagogo (ID31)

Sebastian Basso (ID31)

Mark Jenner (ID31)

Lisa Knight (ID31)

Gavin Fox (BM16)

The ID31 team

## IUCr

Yuji Ohashi

Henk Schenk

## Univ. of Manchester

John Helliwell

## APS

Bob Von Dreele

## AFMB- Marseille

Nicolas Papageorgiou

Bruno Canard

## EMBL - Hamburg

Matthias Wilmanns

Nikos Pinotsis

## EPFL

Marc Schiltz

Celine Besnard

## SPring8

Keiko Miura

## CRMCN, Marseille

Françoise Bonneté

Marion Giffard

## IBS - Grenoble

Richard Kahn

## Novo Nordisk

Gerd Schluckebier

Mathias Norrman

## Sanofi - Aventis

Bertrand Castro

M. El Hajji

## CCP4 & CCP14

PANalytical  
BRUKER